**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/31, 15/62, C07K 14/315, A61K 39/09 — A2

(11) International Publication Number: **WO 00/39299**

(43) International Publication Date: 6 July 2000 (06.07.00)

(21) International Application Number: PCT/CA99/01218

(22) International Filing Date: 20 December 1999 (20.12.99)

(30) Priority Data:
60/113,800 23 December 1998 (23.12.98) US

(71) Applicant *(for all designated States except US)*: BIOCHEM PHARMA INC. [CA/CA]; 275 Armand Frappier Boulevard, Laval, Québec H7V 4A7 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HAMEL, Josée [CA/CA]; (CA). BRODEUR, Bernard, R. [CA/CA]; 2401 Maritain, Sillery, Québec G1T 1N6 (CA). PINEAU, Isabelle [CA/CA]; 3035 Rue des Châtelets, Sainte–Foy, Québec G1V 3Y7 (CA). MARTIN, Denis [CA/CA]; 4728–G Gaboury, St–Augustin–DesMaures, Québec G3A 1E9 (CA). RIOUX, Clément [CA/CA]; 12 Croissant Barabé, Ile Bizard, Québec H9E 1J1 (CA). CHARLAND, Nathalie [CA/CA]; 4340 Rue du Rapide, Apt. 8, Charny, Québec J6X 3N6 (CA).

(74) Agents: MURPHY, Kevin, P. et al.; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montreal, Québec H3A 2Y3 (CA).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: NOVEL STREPTOCOCCUS ANTIGENS

| BVH11-2 SP64 | BVH11 SP63 | BVH11 JNR.7/87 | BVH11-2 JNR.7/87 | BVH11 WU2 | BVH11-2 WU2 | BVH11 A66 | BVH11-2 A66 | BVH11 P4241 | BVH11-2 P4241 | BVH11 Rx-1 | BVH11-2 Rx-1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I 81% S 86% | I 88% S 90% | I 88% S 91% | I 82% S 87% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 88% S 91% | I 81% S 85% | BVH11 SP64 |
| | I 87% S 90% | I 87% S 90% | I 98% S 98% | I 95% S 96% | I 96% S 97% | I 95% S 96% | I 96% S 97% | I 95% S 96% | I 96% S 97% | I 87% S 90% | I 94% S 95% | BVH11-2 SP64 |
| | | I 96% S 96% | I 88% S 91% | I 88% S 91% | I 87% S 90% | I 88% S 91% | I 87% S 90% | I 88% S 91% | I 87% S 90% | I 97% S 97% | I 89% S 91% | BVH11 SP63 |
| | | | I 87% S 90% | I 87% S 91% | I 86% S 90% | I 87% S 91% | I 86% S 90% | I 87% S 91% | I 86% S 90% | I 96% S 96% | I 88% S 90% | BVH11 JNR.7/87 |
| | | | | I 96% S 97% | I 97% S 98% | I 96% S 97% | I 97% S 98% | I 96% S 97% | I 97% S 98% | I 87% S 90% | I 94% S 95% | BVH11-2 JNR.7/87 |
| | | | | | I 98% S 98% | I 92% S 94% | I 98% S 98% | I 99% S 99% | I 98% S 98% | I 87% S 91% | I 92% S 94% | BVH11 WU2 |
| | | | | | | I 98% S 98% | I 99% S 99% | I 98% S 98% | I 99% S 99% | I 86% S 90% | I 93% S 95% | BVH11-2 WU2 |
| | | | | | | | I 99% S 99% | I 100% | I 99% S 99% | I 87% S 91% | I 92% S 94% | BVH11 A66 |
| | | | | | | | | I 99% S 99% | I 99% S 99% | I 86% S 90% | I 93% S 95% | BVH11-2 A66 |
| | | | | | | | | | I 99% S 99% | I 87% S 91% | I 92% S 94% | BVH11 P4241 |
| | | | | | | | | | | I 86% S 90% | I 93% S 95% | BVH11-2 P4241 |
| | | | | | | | | | | | I 91% S 92% | BVH11 Rx-1 |

(57) Abstract

Streptococcus proteins and polynucleotides encoding them are disclosed. Said proteins are antigenic and therefore useful vaccine components for the prophylaxis or therapy of streptococcus infection in animals. Also disclosed are recombinant methods of producing the protein antigens as well as diagnostic assays for detecting streptococcus bacterial infection.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | |
|---|---|
| AL | Albania |
| AM | Armenia |
| AT | Austria |
| AU | Australia |
| AZ | Azerbaijan |
| BA | Bosnia and Herzegovina |
| BB | Barbados |
| BE | Belgium |
| BF | Burkina Paso |
| BG | Bulgaria |
| BJ | Benin |
| BR | Brazil |
| BY | Belarus |
| CA | Canada |
| CF | Central African Republic |
| CG | Congo |
| CH | Switzerland |
| CI | Côte d'Ivoire |
| CM | Cameroon |
| CN | China |
| CU | Cuba |
| CZ | Czech Republic |
| DE | Germany |
| DK | Denmark |
| EE | Estonia |
| ES | Spain |
| FI | Finland |
| FR | France |
| GA | Gabon |
| GB | United Kingdom |
| GE | Georgia |
| GH | Ghana |
| GN | Guinea |
| GR | Greece |
| HU | Hungary |
| IE | Ireland |
| IL | Israel |
| IS | Iceland |
| IT | Italy |
| JP | Japan |
| KE | Kenya |
| KG | Kyrgyzstan |
| KP | Democratic People's Republic of Korea |
| KR | Republic of Korea |
| KZ | Kazakstan |
| LC | Saint Lucia |
| LI | Liechtenstein |
| LK | Sri Lanka |
| LR | Liberia |
| LS | Lesotho |
| LT | Lithuania |
| LU | Luxembourg |
| LV | Latvia |
| MC | Monaco |
| MD | Republic of Moldova |
| MG | Madagascar |
| MK | The former Yugoslav Republic of Macedonia |
| ML | Mali |
| MN | Mongolia |
| MR | Mauritania |
| MW | Malawi |
| MX | Mexico |
| NE | Niger |
| NL | Netherlands |
| NO | Norway |
| NZ | New Zealand |
| PL | Poland |
| PT | Portugal |
| RO | Romania |
| RU | Russian Federation |
| SD | Sudan |
| SE | Sweden |
| SG | Singapore |
| SI | Slovenia |
| SK | Slovakia |
| SN | Senegal |
| SZ | Swaziland |
| TD | Chad |
| TG | Togo |
| TJ | Tajikistan |
| TM | Turkmenistan |
| TR | Turkey |
| TT | Trinidad and Tobago |
| UA | Ukraine |
| UG | Uganda |
| US | United States of America |
| UZ | Uzbekistan |
| VN | Viet Nam |
| YU | Yugoslavia |
| ZW | Zimbabwe |

# NOVEL STREPTOCOCCUS ANTIGENS

## FIELD OF THE INVENTION

The present invention is related to antigens, more particularly protein antigens of streptococcus pneumoniaepathogen which are useful as vaccine components for therapy and/or prophylaxis.

## BACKGROUND OF THE INVENTION

S. pneumoniae is an important agent of disease in man especially among infants, the elderly and immunocompromised persons. It is a bacterium frequently isolated from patients with invasive diseases such as bacteraemia/septicaemia, pneumonia, meningitis with high morbidity and mortality throughout the world. Even with appropriate antibiotic therapy, pneumococcal infections still result in many deaths. Although the advent of antimicrobial drugs has reduced the overall mortality from pneumococcal disease, the presence of resistant pneumococcal organisms has become a major problem in the world today. Effective pneumococcal vaccines could have a major impact on the morbidity and mortality associated with S. pneumoniae disease. Such vaccines would also potentially be useful to prevent otitis media in infants and young children.

Efforts to develop a pneumococcal vaccine have generally concentrated on generating immune responses to the pneumococcal capsular polysaccharide. More than 80 pneumococcal capsular serotypes have been identified on the basis of antigenic differences. The currently available pneumococcal vaccine, comprising 23 capsular polysaccharides

that most frequently caused disease, has significant shortcomings related primarily to the poor immunogenicity of some capsular polysaccharides, the diversity of the serotypes and the differences in the distribution of serotypes over time, geographic areas and age groups. In particular, the failure of existing vaccines and capsular conjugate vaccines currently in development to protect young children against all serotypes spurres evaluation of other S. pneumoniae components. Although immunogenicity of capsular polysaccharides can be improved, serotype specificity will still represent a major limitation of polysaccharide-based vaccines. The use of a antigenically conserved immunogenic pneumococcal protein antigen, either by itself or in combination with additional components, offers the possibility of a protein-based pneumococcal vaccine.

PCT Publication number WO98/18930 published may 7 1998 entitled "*Streptococcus Pneumoniae* antigens and vaccines" describes certain polypeptides which are claimed to be antigenic. However, no biological activity of these polypeptides is reported.

Therefore their remains an unmet need for Streptococcus antigens that may be used as vaccine components for the prophylaxis and/or therapy of Streptococcus infection.

## SUMMARY OF THE INVENTION

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55**

**to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

In other aspects, there are provided vectors comprising polynucleotides of the invention operably linked to an expression control region, as well as host cells transfected with said vectors and methods of producing polypeptides comprising culturing said host cells under conditions suitable for expression.

In yet another aspect, there are provided novel polypeptides encoded by polynucleotides of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is the DNA sequence of BVH-3 gene; **SEQ ID NO: 1.**

Figure 2 is the amino acid sequence of BVH-3 protein; **SEQ ID NO: 2.**

Figure 3 is the DNA sequence of BVH-11 gene; **SEQ ID NO: 3.**

Figure 4 is the amino acid sequence of BVH-11 protein; **SEQ ID NO: 4.**

Figure 5 is the DNA sequence of BVH-28 gene; **SEQ ID NO: 5.**

Figure 6 is the amino acid sequence of BVH-28 protein; **SEQ ID NO: 6.**

Figure 7 is the DNA sequence of BVH-3A gene which corresponds to the 5' terminal end of BVH-3; **SEQ ID NO: 7.**

Figure 8 is the amino acid sequence of BVH-3A protein; **SEQ ID NO: 8.**

Figure 9 is the DNA sequence of BVH-3B gene which corresponds to the 3' terminal end of BVH-3; **SEQ ID NO: 9.**

Figure 10 is the amino acid sequence of BVH-3B protein; **SEQ ID NO: 10.**

Figure 11 depicts the comparison of the predicted amino acid sequences of the BVH-3 open reading frames from WU2, RX1, JNR.7/87, SP64, P4241 and A66 S. pneumoniae strains by using the program Clustal W from MacVector sequence analysis software (version 6.5). Underneath the alignment, there is a consensus line where * and . characters indicate identical and similar amino acid residues, respectively.

Figure 12 depicts the comparison of the predicted amino acid sequences of the BVH-11 open reading frames from WU2, Rx1, JNR.7/87, SP64, P4241, A66 and SP63 S. pneumoniae strains by using the program Clustal W from MacVector sequence analysis software (version 6.5). Underneath the alignment, there is a consensus line where * and . characters indicate identical and similar amino acid residues, respectively.

Figure 13 depicts the comparison of the predicted amino acid sequences of the BVH-11 proteins from various S. pneumoniae strains. The degrees of identity (I) and similarity (S) were determined by using the program Clustal W from MacVector sequence analysis software (version 6.5).

Figure 14 is a DNA sequence containing the complete BVH-3 gene (open reading frame "ORF" at nucleotides 1777 to 4896); **SEQ ID NO: 11.**

Figure 15 is a DNA sequence containing the complete BVH-11 gene (ORF at nucleotides 45 to 2567); **SEQ ID NO: 12.**

Figure 16 is a DNA sequence containing the complete BVH-11-2 gene (ORF at nucleotides 114 to 2630); **SEQ ID NO: 13.**

Figure 17 is the amino acid sequence of BVH-11-2 protein; **SEQ ID NO: 14.**

Figure 18 is the DNA sequence of SP63 BVH-3 gene; **SEQ ID NO:15.**

Figure 19 is the amino acid sequence of SP63 BVH-3 protein; **SEQ ID NO: 16.**

Figure 20 is the amino acid sequence of BVH-3M protein; **SEQ ID NO: 55.**

Figure 21 is the amino acid sequence of BVH-3AD protein; **SEQ ID NO: 56.**

Figure 22 is the amino acid sequence of L-BVH-3-AD protein; **SEQ ID NO: 57.**

Figure 23 is the amino acid sequence of NEW12 protein; **SEQ ID NO: 58.**

Figure 24 is the amino acid sequence of BVH-3C protein; **SEQ ID NO: 59.**

Figure 25 is the amino acid sequence of BVH-11M protein; **SEQ ID NO: 60.**

Figure 26 is the amino acid sequence of BVH-11A protein; **SEQ ID NO: 61.**

Figure 27 is the amino acid sequence of BVH-11B (also called New13) protein; **SEQ ID NO: 62.**

Figure 28 is the amino acid sequence of BVH-11C protein; **SEQ ID NO: 63.**

Figure 29 is the amino acid sequence of NEW1 protein; **SEQ ID NO: 64.**

Figure 30 is the amino acid sequence of NEW2 protein; **SEQ ID NO: 65.**

Figure 31 is the amino acid sequence of NEW3 protein; **SEQ ID NO: 66.**

Figure 32 is the amino acid sequence of NEW4 protein; **SEQ ID NO: 67.**

Figure 33 is the amino acid sequence of NEW5 protein; **SEQ ID NO: 68.**

Figure 34 is the amino acid sequence of NEW6 protein; **SEQ ID NO: 69.**

Figure 35 is the amino acid sequence of NEW7 protein; **SEQ ID NO: 70.**

Figure 36 is the amino acid sequence of NEW8 protein; **SEQ ID NO: 71.**

Figure 37 is the amino acid sequence of NEW9 protein; **SEQ ID NO: 72.**

Figure 38 is the amino acid sequence of BVH-11-2M protein; **SEQ ID NO: 73.**

Figure 39 is the amino acid sequence of NEW10 protein; **SEQ ID NO: 74.**

Figure 40 is the amino acid sequence of NEW11 protein; **SEQ ID NO: 75.**

Figure 41 is the DNA sequence of NEW12 gene; **SEQ ID NO: 76.**

Figure 42 is the amino acid sequence of NEW14 protein; **SEQ ID NO: 77.**

Figure 43 is the amino acid sequence of NEW15 protein; **SEQ ID NO: 78.**

Figure 44 is the amino acid sequence of NEW16 protein; **SEQ ID NO: 79.**

Figure 45 is the DNA sequence of GBS BVH-71 gene; **SEQ ID NO: 80.**

Figure 46 is the amino acid sequence of GBS BVH-71 protein; **SEQ ID NO: 81.**

Figure 47 is the DNA sequence of GAS BVH-71 gene; **SEQ ID NO:82.**

Figure 48 is the amino acid sequence of GAS BVH-71 protein; **SEQ ID NO:83.**

## DETAILED DESCRIPTION OF THE INVENTION

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55**

**to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 95% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 8, 10, 14, 16, 55 to 75, 77 to 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 8, 10, 16, 55, 56, 57, 58, 59, 64, 65, 66, 78** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 8, 10, 16, 55, 56, 57, 59, 64, 65, 66, 78** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 4, 14, 58, 60, 61, 62, 63, 67, 68, 69, 70, 71, 72, 73, 74, 75, 77, 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 4, 14, 60, 61, 62, 63, 67, 68, 69, 70, 71, 72, 73, 74, 75, 77, 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 10, 14, 16, 55 to 75, 77 to 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence chosen from **SEQ ID NOs: 10, 55 to 75, 77, 78, 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an

isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence chosen from **SEQ ID NOs: 55 to 75, 77, 78, 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 10, 14, 16** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising a sequence chosen from **SEQ ID NOs: 2, 4, 14, 16** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 2** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 4** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 10** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 14** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 16** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 58** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 60** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 62** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 64** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 67** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 68** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 69** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 72** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 74** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention provides an isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide comprising sequence **SEQ ID NO: 77** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4 , 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 8, 10, 14, 16, 55 to 75, 77 to 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 10, 14, 16, 55 to 75, 77 to 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 10, 14, 16** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 2** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 4** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 10** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 14** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 16** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 10, 55 to 75, 77, 78, 79** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NO: 10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NO: 10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NO: 10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence chosen from **SEQ ID NO: 10, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 58** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 62** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 64** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to

polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 67** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 68** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 74** or fragments, analogs or derivatives thereof.

According to one aspect, the present invention relates to polypeptides characterized by the amino acid sequence comprising sequence **SEQ ID NO: 77** or fragments, analogs or derivatives thereof.

In a further embodiment, the present invention also relates to chimeric polypeptides which comprise one or more polypeptides or fragments, analogs or derivatives thereof as described in the present application.

In a further embodiment, the present invention also relates to chimeric polypeptides which comprise one or more polypeptides or fragments, analogs or derivatives thereof as defined in the figures of the present application.

In a further embodiment, the present application also relates to chimeric polypeptides which comprise two or more polypeptides chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof ;provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

In a further embodiment, the chimeric polypeptide will comprise two or more polypeptides chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof; provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

In a further embodiment, the chimeric polypeptide will comprise two or more polypeptides chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof; provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

In a further embodiment, the chimeric polypeptide will comprise two or more polypeptides chosen from **SEQ ID NOs :10, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof; provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

In a further embodiment, the chimeric polypeptide will comprise between 2 and 5 polypeptides.

In a further embodiment, the chimeric polypeptide will comprise between 2 and 4 polypeptides.

In a further embodiment, the chimeric polypeptide will comprise between 2 and 3 polypeptides.

In a further embodiment, the chimeric polypeptide will comprise 2 polypeptides.

In a further embodiment, there is provided a chimeric polypeptide of formula (I):

**A-(B)$_m$-(C)$_n$-D** (I)

Wherein;

**m** is 0 or 1,

**n** is 0 or 1,

**A** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof;

**B** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof;

**C** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof; and

**D** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

In a further embodiment,

**A** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof;

**B** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77**, or fragments, analogs or derivatives thereof;

**C** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof; and

**D** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof.

In a further embodiment,

**A** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof;
**B** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 74, 77,** or fragments, analogs or derivatives thereof;
**C** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof; and
**D** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.

In one embodiment, chimeric polypeptides of the present invention comprise those wherein the following embodiments are present, either independently or in combination.

In a further embodiment, **A** is **SEQ ID NOs :10, 58, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :10** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :58** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :62** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :64** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :67** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :68** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :74** or fragments, analogs or derivatives thereof.
In a further embodiment, **A** is **SEQ ID NO :77** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NOs :10, 58, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :10** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :58** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :64** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :64** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :67** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :68** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO :74** or fragments, analogs or derivatives thereof.

In a further embodiment, **B** is **SEQ ID NO : 77** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NOs :10, 58, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO :10** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO :58** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO : 62** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO :64** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO : 67** or fragments, analogs or derivatives thereof.

In a further embodiment, **C** is **SEQ ID NO : 68** or fragments,

analogs or derivatives thereof.
In a further embodiment, **C** is **SEQ ID NO : 74** or fragments, analogs or derivatives thereof.
In a further embodiment, **C** is **SEQ ID NO : 77** or fragments, analogs or derivatives thereof.

In a further embodiment, **D** is **SEQ ID NO :10, 58, 62, 64, 67, 68, 74, 77** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :10** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :58** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :62** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :64** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :67** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :68** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :74** or fragments, analogs or derivatives thereof.
In a further embodiment, **D** is **SEQ ID NO :77** or fragments, analogs or derivatives thereof.

In a further embodiment, **m** is 0.
In a further embodiment, **n** is 0.

In a further embodiment, **m** and **n** are 0.

In a further embodiment, **m** and **n** are 0, **A** is **SEQ ID NO:64** or fragments, analogs or derivatives thereof, **B** is **SEQ ID**

**NO:62** or fragments, analogs or derivatives thereof. In a further embodiment, **m** and **n** are 0, **A** is **SEQ ID NO:62** or fragments, analogs or derivatives thereof, **B** is **SEQ ID NO:64** or fragments, analogs or derivatives thereof.

In accordance with the present invention, all nucleotides encoding polypeptides and chimeric polypeptides are within the scope of the present invention.

In a further embodiment, the polypeptides or chimeric polypeptides in accordance with the present invention are antigenic.

In a further embodiment, the polypeptides or chimeric polypeptides in accordance with the present invention can elicit an immune response in an individual.

In a further embodiment, the present invention also relates to polypeptides which are able to raise antibodies having binding specificity to the polypeptides or chimeric polypeptides of the present invention as defined above.

An antibody that " has binding specificity" is an antibody that recognizes and binds the selected polypeptide but which does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which naturally includes the selected peptide. Specific binding can be measured using an ELISA assay in which the selected polypeptide is used as an antigen.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In

case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

As used herein, "fragments", "derivatives" or "analogs" of the polypeptides of the invention include those polypeptides in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably conserved) and which may be natural or unnatural. In one embodiment, derivatives and analogs of polypeptides of the invention will have about 70% identity with those sequences illustrated in the figures or fragments thereof. That is, 70% of the residues are the same. In a further embodiment, polypeptides will have greater than 75% homology. In a further embodiment, polypeptides will have greater than 80% homology. In a further embodiment, polypeptides will have greater than 85% homology. In a further embodiment, polypeptides will have greater than 90% homology. In a further embodiment, polypeptides will have greater than 95% homology. In a further embodiment, polypeptides will have greater than 99% homology. In a further embodiment, derivatives and analogs of polypeptides of the invention will have fewer than about 20 amino acid residue substitutions, modifications or deletions and more preferably less than 10. Preferred substitutions are those known in the art as conserved i.e. the substituted residues share physical or chemical properties such as hydrophobicity, size, charge or functional groups.

In accordance with the present invention, polypeptides of the invention include both polypeptides and chimeric polypeptides.

Also included are polypeptides which have fused thereto

other compounds which alter the polypeptides biological or pharmacological properties i.e. polyethylene glycol (PEG) to increase half-life; leader or secretory amino acid sequences for ease of purification; prepro- and pro- sequences; and (poly)saccharides.

Furthermore, in those situations where amino acid regions are found to be polymorphic, it may be desirable to vary one or more particular amino acids to more effectively mimic the different epitopes of the different streptococcus strains.

Moreover, the polypeptides of the present invention can be modified by terminal -$NH_2$ acylation (eg. by acetylation, or thioglycolic acid amidation, terminal carbosy amidation, e.g. with ammonia or methylamine) to provide stability, increased hydrophobicity for linking or binding to a support or other molecule.

Also contemplated are hetero and homo polypeptide multimers of the polypeptide fragments, analogues and derivatives. These polymeric forms include, for example, one or more polypeptides that have been cross-linked with cross-linkers such as avidin/biotin, gluteraldehyde or dimethyl-superimidate. Such polymeric forms also include polypeptides containing two or more tandem or inverted contiguous sequences, produced from multicistronic mRNAs generated by recombinant DNA technology.
Preferably, a fragment, analog or derivative of a polypeptide of the invention will comprise at least one antigenic region i.e. at least one epitope.

In order to achieve the formation of antigenic polymers (i.e. synthetic multimers), polypeptides may be utilized having bishaloacetyl groups, nitroarylhalides, or the like, where the reagents being specific for thio groups.

Therefore, the link between two mercapto groups of the different peptides may be a single bond or may be composed of a linking group of at least two, typically at least four, and not more than 16, but usually not more than about 14 carbon atoms.

In a particular embodiment, polypeptide fragments, analogs and derivatives of the invention do not contain a methionine (Met) starting residue. Preferably, polypeptides will not incorporate a leader or secretory sequence (signal sequence). The signal portion of a polypeptide of the invention may be determined according to established molecular biological techniques. In general, the polypeptide of interest may be isolated from a streptococcus culture and subsequently sequenced to determine the initial residue of the mature protein and therefore the sequence of the mature polypeptide.

According to another aspect, there are provided vaccine compositions comprising one or more streptococcus polypeptides of the invention in admixture with a pharmaceutically acceptable carrier diluent or adjuvant. Suitable adjuvants include oils i.e. Freund's complete or incomplete adjuvant; salts i.e. $AlK(SO_4)_2$, $AlNa(SO_4)_2$, $AlNH_4(SO_4)_2$, silica, kaolin, carbon polynucleotides i.e. poly IC and poly AU. Preferred adjuvants include QuilA and Alhydrogel. Vaccines of the invention may be administered parenterally by injection, rapid infusion, nasopharyngeal absorption, dermoabsorption, or bucal or oral. Pharmaceutically acceptable carriers also include tetanus toxoid.

Vaccine compositions of the invention are used for the treatment or prophylaxis of streptococcus infection and/or diseases and symptoms mediated by streptococcus infection as described in P.R. Murray (Ed, in chief),E.J. Baron, M.A.

Pfaller, F.C. Tenover and R.H. Yolken. Manual of Clinical Microbiology, ASM Press, Washington, D.C. sixth edition, 1995, 1482p which are herein incorporated by reference. In one embodiment, vaccine compositions of the present invention are used for the treatment or prophylaxis of meningitis, otitis media, bacteremia or pneumonia. In one embodiment, vaccine compositions of the invention are used for the treatment or prophylaxis of *streptococcus* infection and/or diseases and symptoms mediated by *streptococcus* infection, in particular S.pneumoniae, group A *streptococcus* (*pyogenes*), group B *streptococcus* (GBS or *agalactiae*), *dysgalactiae*, *uberis*, *nocardia* as well as *Staphylococcus aureus*. In a further embodiment, the streptococcus infection is S.pneumoniae.

In a particular embodiment, vaccines are administered to those individuals at risk of streptococcus infection such as infants, elderly and immunocompromised individuals.

As used in the present application, the term " individuals" include mammals. In a further embodiment, the mammal is human.

Vaccine compositions are preferably in unit dosage form of about 0.001 to 100 µg/kg (antigen/body weight) and more preferably 0.01 to 10 µg/kg and most preferably 0.1 to 1 µg/kg 1 to 3 times with an interval of about 1 to 6 week intervals between immunizations.

According to another aspect, there are provided polynucleotides encoding polypeptides characterized by the amino acid sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

In one embodiment, polynucleotides are those illustrated in **SEQ ID Nos: 1, 3, 5, 7, 9, 11, 12, 13, 15, 76, 80, 82** which may include the open reading frames (ORF), encoding polypeptides of the invention. It will be appreciated that the polynucleotide sequences illustrated in the figures may be altered with degenerate codons yet still encode the polypeptides of the invention. Accordingly the present invention further provides polynucleotides which hybridize to the polynucleotide sequences herein above described (or the complement sequences thereof) having 50% identity between sequences. In one embodiment, at least 70% identity between sequences. In one embodiment, at least 75% identity between sequences. In one embodiment, at least 80% identity between sequences. In one embodiment, at least 85% identity between sequences. In one embodiment, at least 90% identity between sequences. In a further embodiment, polynucleotides are hybridizable under stringent conditions i.e. having at least 95% identity. In a further embodiment, more than 97% identity.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 3, 7, 9, 11, 12, 13, 15, 76, 80, 82** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 3, 9, 11, 12, 13, 15, 76, 80, 82** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 3, 9, 11, 12, 13, 15, 76** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those

illustrated in **SEQ ID NOs : 1, 3, 7, 9, 11, 12, 13, 15, 76** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 7, 9, 11, 15, 76** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 9, 11, 15, 76** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 1, 7, 9, 11** which may include the open reading frames (ORF), encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO : 1,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :7,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :9,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :11,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :15,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NOs : 3, 12, 13, 76,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :3,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :12,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :13,** encoding polypeptides of the invention.

In a further embodiment, polynucleotides are those illustrated in **SEQ ID NO :76,** encoding polypeptides of the invention.

As will be readily appreciated by one skilled in the art, polynucleotides include both DNA and RNA.

The present invention also includes polynucleotides complementary to the polynucleotides described in the present application.

In a further aspect, polynucleotides encoding polypeptides of the invention, or fragments, analogs or derivatives thereof, may be used in a DNA immunization method. That is, they can be incorporated into a vector which is

replicable and expressible upon injection thereby producing the antigenic polypeptide in vivo. For example polynucleotides may be incorporated into a plasmid vector under the control of the CMV promoter which is functional in eukaryotic cells. Preferably the vector is injected intramuscularly.

According to another aspect, there is provided a process for producing polypeptides of the invention by recombinant techniques by expressing a polynucleotide encoding said polypeptide in a host cell and recovering the expressed polypeptide product. Alternatively, the polypeptides can be produced according to established synthetic chemical techniques i.e. solution phase or solid phase synthesis of oligopeptides which are ligated to produce the full polypeptide (block ligation).

General methods for obtention and evaluation of polynucleotides and polypeptides are described in the following references: Sambrook et al, Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor, N.Y., 1989; Current Protocols in Molecular Biology, Edited by Ausubel F.M. et al., John Wiley and Sons, Inc. New York; PCR Cloning Protocols, from Molecular Cloning to Genetic Engineering, Edited by White B.A., Humana Press, Totowa, New Jersey, 1997, 490 pages; Protein Purification, Principles and Practices, Scopes R.K., Springer-Verlag, New York, 3rd Edition, 1993, 380 pages; Current Protocols in Immunology, Edited by Coligan J.E. et al., John Wiley & Sons Inc., New York which are herein incorporated by reference.

For recombinant production, host cells are transfected with vectors which encode the polypeptide, and then cultured in a nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes.

Suitable vectors are those that are viable and replicable in the chosen host and include chromosomal, non-chromosomal and synthetic DNA sequences e.g. bacterial plasmids, phage DNA, baculovirus, yeast plasmids, vectors derived from combinations of plasmids and phage DNA. The polypeptide sequence may be incorporated in the vector at the appropriate site using restriction enzymes such that it is operably linked to an expression control region comprising a promoter, ribosome binding site (consensus region or Shine-Dalgarno sequence), and optionally an operator (control element). One can select individual components of the expression control region that are appropriate for a given host and vector according to established molecular biology principles (Sambrook et al, Molecular Cloning: A Laboratory Manual, 2nd ed, Cold Spring Harbor, N.Y., 1989; Current Protocols in Molecular Biology, Edited by Ausubel F.M. et al., John Wiley and Sons, Inc. New York incorporated herein by reference). Suitable promoters include but are not limited to LTR or SV40 promoter, E.coli lac, tac or trp promoters and the phage lambda $P_L$ promoter. Vectors will preferably incorporate an origin of replication as well as selection markers i.e. ampicilin resistance gene. Suitable bacterial vectors include pET, pQE70, pQE60, pQE-9, pbs, pD10 phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A, ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 and eukaryotic vectors pBlueBacIII, pWLNEO, pSV2CAT, pOG44, pXT1, pSG, pSVK3, pBPV, pMSG and pSVL. Host cells may be bacterial i.e. E.coli, Bacillus subtilis, Streptomyces; fungal i.e. Aspergillus niger, Aspergillus nidulins; yeast i.e. Saccharomyces or eukaryotic i.e. CHO, COS.

Upon expression of the polypeptide in culture, cells are typically harvested by centrifugation then disrupted by physical or chemical means (if the expressed polypeptide is not secreted into the media) and the resulting crude

extract retained to isolate the polypeptide of interest. Purification of the polypeptide from culture media or lysate may be achieved by established techniques depending on the properties of the polypeptide i.e. using ammonium sulfate or ethanol precipitation , acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, hydroxylapatite chromatography and lectin chromatography. Final purification may be achieved using HPLC.

The polypeptide may be expressed with or without a leader or secretion sequence. In the former case the leader may be removed using post-translational processing (see US 4,431,739; US 4,425,437; and US 4,338,397 incorporated herein by reference) or be chemically removed subsequent to purifying the expressed polypeptide.

According to a further aspect, the streptococcus polypeptides of the invention may be used in a diagnostic test for streptococcus infection, in particular S. pneumoniae infection. Several diagnostic methods are possible, for example detecting streptococcus organism in a biological sample, the following procedure may be followed:

a) obtaining a biological sample from a patient;

b) incubating an antibody or fragment thereof reactive with a streptococcus polypeptide of the invention with the biological sample to form a mixture; and

c) detecting specifically bound antibody or bound fragment in the mixture which indicates the presence of streptococcus.

Alternatively, a method for the detection of antibody specific to a streptococcus antigen in a biological sample containing or suspected of containing said antibody may be performed as follows:

a) obtaining a biological sample from a patient;

b) incubating one or more streptococcus polypeptides of the invention or fragments thereof with the biological sample to form a mixture; and

c) detecting specifically bound antigen or bound fragment in the mixture which indicates the presence of antibody specific to streptococcus.

One of skill in the art will recognize that this diagnostic test may take several forms, including an immunological test such as an enzyme-linked immunosorbent assay (ELISA), a radioimmunoassay or a latex agglutination assay, essentially to determine whether antibodies specific for the protein are present in an organism.

The DNA sequences encoding polypeptides of the invention may also be used to design DNA probes for use in detecting the presence of streptococcus in a biological sample suspected of containing such bacteria. The detection method of this invention comprises:

a) obtaining the biological sample from a patient;

b) incubating one or more DNA probes having a DNA sequence encoding a polypeptide of the invention or fragments thereof with the biological sample to form a mixture; and

c) detecting specifically bound DNA probe in the mixture which indicates the presence of streptococcus bacteria.

The DNA probes of this invention may also be used for detecting circulating streptococcus i.e. S.pneumoniaenucleic acids in a sample, for example using a polymerase chain reaction, as a method of diagnosing streptococcus infections. The probe may be synthesized using conventional techniques and may be immobilized on a solid phase, or may be labelled with a detectable label. A preferred DNA probe for this application is an oligomer

having a sequence complementary to at least about 6 contiguous nucleotides of the streptococcus pneumoniae polypeptides of the invention.

Another diagnostic method for the detection of streptococcus in a patient comprises:

a) labelling an antibody reactive with a polypeptide of the invention or fragment thereof with a detectable label;

b) administering the labelled antibody or labelled fragment to the patient; and

c) detecting specifically bound labelled antibody or labelled fragment in the patient which indicates the presence of streptococcus.

A further aspect of the invention is the use of the streptococcus polypeptides of the invention as immunogens for the production of specific antibodies for the diagnosis and in particular the treatment of streptococcus infection. Suitable antibodies may be determined using appropriate screening methods, for example by measuring the ability of a particular antibody to passively protect against streptococcus infection in a test model. One example of an animal model is the mouse model described in the examples herein. The antibody may be a whole antibody or an antigen-binding fragment thereof and may belong to any immunoglobulin class. The antibody or fragment may be of animal origin, specifically of mammalian origin and more specifically of murine, rat or human origin. It may be a natural antibody or a fragment thereof, or if desired, a recombinant antibody or antibody fragment. The term recombinant antibody or antibody fragment means antibody or antibody fragment which was produced using molecular biology techniques. The antibody or antibody fragments may be polyclonal, or preferably monoclonal. It may be specific for a number of epitopes associated with the

streptococcus pneumoniae polypeptides but is preferably specific for one.

Without limiting its scope, the present invention also relates to new antigens designated BVH-3, BVH-11, BVH-11-2, BVH-28 and BVH-71. The present invention also relates to truncated polypeptides comprising fragments of the new antigens designated BVH-3, BVH-11, BVH-11-2, BVH-28 and BVH-71. The present invention also relates to chimeric polypeptides comprising fragments of the new antigens designated BVH-3, BVH-11, BVH-11-2, BVH-28 and BVH-71. The following is a reference table summarizing the relation between the antigens of the present invention:

| **Family** | **Nucleotide SEQ ID NO** | **Polypeptide SEQ ID NO** |
|---|---|---|
| BVH-3 | | |
| BVH-3 | 1, 11 | 2 |
| BVH-3A | 7 | 8 |
| BVH-3B | 9 | 10 |
| BVH-3 SP63 | 15 | 16 |
| BVH-3M | | 55 |
| BVH-3AD | | 56 |
| L-BVH-3AD | | 57 |
| New12 | 76 | 58 |
| BVH-3C | | 59 |
| New1 | | 64 |
| New2 | | 65 |
| New3 | | 66 |
| New15 | | 78 |
| BVH-11 | | |
| BVH-11 | 3, 12 | 4 |
| BVH-11-2 | 13 | 14 |
| BVH-11M | | 60 |
| BVH-11A | | 61 |
| BVH-11B also referred to as NEW13 | | 62 |
| BVH-11C | | 63 |
| New4 | | 67 |
| New5 | | 68 |

| Family | Nucleotide SEQ ID NO | Polypeptide SEQ ID NO |
|---|---|---|
| New6 | | 69 |
| New7 | | 70 |
| New8 | | 71 |
| New9 | | 72 |
| BVH-11-2M | | 73 |
| New10 | | 74 |
| New11 | | 75 |
| New12 | 76 | 58 |
| New14 | | 77 |
| New16 | | 79 |
| BVH-28 | | |
| BVH-28 | 5 | 6 |
| BVH-71 | | |
| GBS | 80 | 81 |
| GAS | 82 | 83 |

EXAMPLE 1

This example illustrates the cloning of S. pneumoniae genes.

The coding region of S. pneumoniae gene BVH-3 (**SEQ ID NO: 1**) and the coding region of S. pneumoniae gene BVH-28 (**SEQ ID NO: 5**) were amplified by PCR (DNA Thermal Cycler GeneAmp PCR system 2400 Perkin Elmer, San Jose, CA) from genomic DNA of serogroup 6 S. pneumoniae strain SP64 using the oligos that contained base extensions for the addition of restriction sites BglII (AGATCT) and XbaI (TCTAGA). PCR products were purified from agarose gel using a QIAquick gel extraction kit from QIAgen (Chatsworth, CA), digested BglII-XbaI (Pharmacia Canada Inc, Baie d'Urfé, Canada), extracted with phenol : chloroform and precipitated with ethanol. The Superlinker vector pSL301 (Invitrogen, San Diego, CA) was digested with BglII and XbaI and purified from agarose gel using a QIAquick gel extraction kit from QIAgen (Chatsworth, CA). The BglII-XbaI genomic DNA fragments were ligated to

the BglII-XbaI pSL301 vector. The ligated products were transformed into E. coli strain DH5a [f80 *lacZ* DM15 *end*A1 *rec*A1 *hsd*R17 ($^{r}K^{-m}K^{+}$) *sup*E44 *thi*-11$^{-}$ *gyr*A96 *rel*A1 D(*lac*ZYA-*argF*)U169] (Gibco BRL, Gaithersburg, MD) according to the method of Simanis (Hanahan, D. DNA Cloning, 1985, D.M. Glover (ed), pp. 109-135). Recombinant pSL301 plasmids (rpSL301) containing either BVH-3 or BVH-28 gene were purified using a QIAgen kit (Chatsworth, CA) and DNA inserts were confirmed by nucleotide sequence analysis (Taq Dye Deoxy Terminator Cycle Sequencing kit, ABI, Foster City, CA). Recombinant rpSL301 (rpSL301) were digested with the restriction enzymes BglII (AGATCT) and XhoI (CTCGAG). DNA fragments BglII-XhoI were purified using the QIAquick gel extraction kit from QIAgen (Chatsworth, CA). pET-32c(+) expression vector (Novagen, Madison, WI) containing the thioredoxin-His·Tag sequence was digested with BamHI (GGATCC) and XhoI and gel extracted using the QIAquick gel extraction kit from QIAgen (Chatsworth, CA). The BglII-XhoI DNA fragments were ligated to the BamHI-XhoI pET-32c(+) vector to create the coding sequence for thioredoxin-His·Tag-BVH-3 or thioredoxin-His·Tag-BVH-28 fusion protein. The ligated products were transformed into E. coli strain DH5a [f80 *lacZ* DM15 *end*A1 *rec*A1 *hsd*R17 ($^{r}K^{-m}K^{+}$) *sup*E44 *thi*-11$^{-}$ *gyr*A96 *rel*A1 D(*lac*ZYA-*argF*)U169] (Gibco BRL, Gaithersburg, MD) according to the method of Simanis (Hanahan, D. DNA Cloning, 1985, D.M. Glover (ed), pp. 109-135). Recombinant pET-32c(+) plasmids were purified using a QIAgen kit (Chatsworth, CA) and the nucleotide sequences at the fusion sites of thioredoxin-His·Tag and DNA insert were verified by DNA sequencing (Taq Dye Deoxy Terminator Cycle Sequencing kit, ABI, Foster City, CA).

EXAMPLE 2

This example illustrates the cloning of S. pneumoniae protein genes in CMV plasmid pCMV-GH.

The DNA coding region of a S. pneumoniae protein was inserted in phase downstream of a human growth hormone (hGH) gene which was under the transcriptional control of the cytomegalavirus (CMV) promotor in the plasmid vector pCMV-GH (Tang et al., Nature, 1992, 356 :152). The CMV promotor is non functional plasmid in E. coli cells but active upon administration of the plasmid in eukaryotic cells. The vector also incorporated the ampicillin resistance gene.

The coding region of BVH-3 gene (**SEQ ID NO: 1**) and BVH-28 gene (**SEQ ID NO: 5**) were obtained from rpSL301 (see example 1) using restriction enzymes BglII (AGATCT) and XbaI (TCTAGA). The digested products were purified from agarose gel using the QIAquick gel extraction kit from QIAgen (Chatsworth, CA). The pCMV-GH vector (Laboratory of Dr. Stephen A. Johnston, Department of Biochemistry, The University of Texas, Dallas, Texas) containing the human growth hormone to create fusion proteins was digested with BglII and XbaI and purified from agarose gel using the QIAquick gel extraction kit from QIAgen (Chatsworth, CA). The BglII-XbaI DNA fragments were ligated to the BglII-XbaI pCMV-GH vector to create the hGH-BVH-3 or hGH-BVH-28 fusion protein under the control of the CMV promoter. The ligated products were transformed into E. coli strain DH5a[f80 *lacZ* DM15 *end*A1 *rec*A1 *hsd*R17 ($^{r}K^{-m}K^{+}$) *sup*E44 *thi*-11$^{-}$ *gyr*A96 *rel*A1 D(*lac*ZYA-*arg*F)U169] (Gibco BRL, Gaithersburg, MD) according

to the method of Simanis (Hanahan, D. DNA Cloning, 1985, D.M. Glover (ed), pp. 109-135). The recombinant pCMV plasmids were purified using a QIAgen kit (QIAgen, Chatsworth, CA).

The coding region of BVH-11 gene (**SEQ ID NO: 3**) was amplified by PCR (DNA Thermal Cycler GeneAmp PCR system 2400 Perkin Elmer, San Jose, CA) from genomic DNA of serogroup 6 S. pneumoniae strain SP64 using the oligos that contained base extensions for the addition of restriction sites BglII (AGATCT) and HindIII (AAGCTT). The PCR product was purified from agarose gel using a QIAquick gel extraction kit from QIAgen (Chatsworth, CA), digested with restriction enzymes (Pharmacia Canada Inc, Baie d'Urfe, Canada), extracted with phenol : chloroform and precipitated with ethanol. The pCMV-GH vector (Laboratory of Dr. Stephen A. Johnston, Department of Biochemistry, The University of Texas, Dallas, Texas) was digested with BglII and HindIII and purified from agarose gel using the QIAquick gel extraction kit from QIAgen (Chatsworth, CA). The BglII-HindIII DNA fragment was ligated to the BglII-HindIII pCMV-GH vector to create the hGH-BVH-11 fusion protein under the control of the CMV promoter. The ligated products were transformed into E. coli strain DH5a[f80 *lac*Z DM15 *end*A1 *rec*A1 *hsd*R17 ($^{r}K^{-m}K^{+}$) *sup*E44 *thi*-11$^{-}$ *gyr*A96 *rel*A1 D(*lac*ZYA-*arg*F)U169] (Gibco BRL, Gaithersburg, MD) according to the method of Simanis (Hanahan, D. DNA Cloning, 1985, D.M. Glover (ed), pp. 109-135). The recombinant pCMV plasmid was purified using a QIAgen kit (Chatsworth, CA) and the nucleotide sequence of the DNA insert was verified by DNA sequencing.

EXAMPLE 3

This example illustrates the use of DNA to elicit an immune response to S. pneumoniae antigens.

A group of 8 female BALB/c mice (Charles River, St-Constant, Québec, Canada) were immunized by intramuscular injection of 50 µl three times at two- or three-week intervals with 100 µg of recombinant pCMV-GH encoding the BVH-3, BVH-11 or the BVH-28 gene in presence of 50 µg of granulocyte-macrophage colony-stimulating factor (GM-CSF)- expressing plasmid pCMV-GH-GM-CSF (Laboratory of Dr. Stephen A. Johnston, Department of Biochemistry, The University of Texas, Dallas, Texas). As control, a group of mice were injected with 100 µg of pCMV-GH in presence of 50 µg of pCMV-GH-GM-CSF. Blood samples were collected from the orbital prior to each immunization and seven days following the third injection and serum antibody responses were determined by ELISA using thioredoxin-His·Tag-S. pneumoniae fusion protein as coating antigen. DNA immunization with recombinant plasmid pCMV-GH encoding the BVH-3, BVH-11 or the BVH-28 S. pneumoniae protein induced antibody reactive against the respective recombinant protein. The reciprocal antibody titers, defined as the highest serum dilution at which the absorbance values were 0.1 above the background values, were above $4x10^3$.

EXAMPLE 4

This example illustrates the production and purification of recombinant S. pneumoniae proteins.

The recombinant pET plasmids containing the BVH-3, BVH-11 or the BVH-28 gene corresponding to the **SEQ ID NO: 1** , **SEQ ID NO: 3** or the **SEQ ID NO: 5** respectively were transformed by electroporation (Gene Pulser II apparatus, BIO-RAD Labs, Mississauga, Canada) into E. coli strain AD494 (DE3) (D*ara*$^-$ *leu*7697 D*lac*X74 D*pho*A *Pvu*II *pho*R D*mal*F3 F'[*lac*$^+$(*lac*I$^q$) *pro*] *trx*B::Kan) (Novagen, Madison, WI). In this strain of E. coli, the T7 promotor controlling expression of the fusion protein is specifically recognized by the T7 RNA polymerase (present on the lDE3 prophage) whose gene is under the control of the lac promotor which is inducible by isopropyl-ß-d-thio-galactopyranoside (IPTG). The transformant AD494(DE3)/rpET was grown at 37°C with agitation at 250 rpm in LB broth (peptone 10g/L, yeast extract 5g/L, NaCl 10g/L) containing 100µg of ampicillin (Sigma-Aldrich Canada Ltd., Oakville, Canada) per ml until the $A_{600}$ reached a value of 0.6. In order to induce the production of the thioredoxin-His·Tag-BVH-3, thioredoxin-His·Tag-BVH-11 or thioredoxin-His·Tag-BVH-28 fusion protein, the cells were incubated for 2 additional hours in the presence of IPTG at a final concentration of 1 mM. Induced cells from a 100 ml culture were pelleted by centrifugation and frozen at -70°C.

The purification of the fusion proteins from the soluble cytoplasmic fraction of IPTG-induced AD494(DE3)/rpET was done by affinity chromatography based on the properties of the His·Tag sequence (6 consecutive histidine residues) to bind to divalent cations ($Ni^{2+}$) immobilized on the His·Bind metal chelation resin. Briefly, the pelleted cells obtained from a 100mL culture induced with IPTG were resuspended in

phosphate-buffered (PBS):500mM NaCl pH7.1, sonicated and spun at 20,000 X g for 20 min to remove debris. The supernatant was filtered (0.22µm pore size membrane) and deposited on a HiTrap® 1mL chelating pre-packed ready-to-use column (Pharmacia Biotech, Baie d'Urfé, Canada). The thioredoxin-His·Tag-S. pneumoniae fusion protein was eluted with 1M imidazole-500mM NaCl-PBS pH7.1. The removal of the salt and imidazole from the sample was done by dialysis against PBS at 4°C. The quantities of fusion protein obtained from the soluble fraction of E. coli was estimated by MicroBCA (Pierce, Rockford, Illinois).

EXAMPLE 5

This example illustrates the protection of mice against fatal pneumococcal infection by immunization.

Groups of 8 female BALB/c mice (Charles River) were immunized subcutaneously three times at three-week intervals with either 25 µg of affinity purified thioredoxin-His·Tag-BVH-3 fusion protein in presence of 15 µg of QuilA adjuvant (Cedarlane Laboratories Ltd, Hornby, Canada) or, as control, with QuilA adjuvant alone in PBS. Blood samples were collected from the orbital sinus on day 1, 22 and 43 prior to each immunization and seven days (day 50) following the third injection. One week later the mice were challenged with approximately $10^6$ CFU of the type 3 S. pneumoniae strain WU2. Samples of the S. pneumoniae challenge inoculum were plated on chocolate agar plates to determine the CFU and to verify the challenge dose. Deaths were recorded for

a period of 14 days and on day 14 post-challenge, the surviving mice were sacrificied and blood samples tested for the presence of S. pneumoniae organisms. The survival data are shown in table 1.

Prechallenge sera were analyzed for the presence of antibodies reactive with S. pneumoniae by standard immunoassays. Elisa and immunoblot analyses indicated that immunization with recombinant S. pneumoniae protein produced in E. coli elicited antibodies reactive with both, recombinant and native pneumococcal protein.

Table 1. Protection mediated by recombinant BVH-3 protein

| Immunogen | No. of mice alive : no. of mice dead<br>14 days post-challenge | Median day of death |
|---|---|---|
| BVH-3 | 8 : 0 | >14 |
| none | 0 : 8 | 1 |

All mice immunized with BVH-3 recombinant protein survived to infection while none of the control mice given adjuvant alone survived. There was a significant difference in survival between the two groups of mice ($P<0.0001$, log rank test for nonparametric analysis of survival curves; $P=0.0002$, Fisher's exact test). All hemocultures from surviving mice were negative at day 14 post-challenge.

EXAMPLE 6

This example describes the cloning of BVH-3 and BVH-11 genes from a variety of S. pneumoniae strains and the molecular conservation of these genes.

Molecular analysis of chromosomal DNA from various S. pneumoniae isolates with DNA probes spanning different regions of BVH-3 or BVH-11 revealed the presence of one BVH-3 gene copy and two BVH-11 gene copies. The two BVH-11 gene copies are not identical and the genes were arbitrarily designated BVH-11 (SEQ ID NO:12; ORF at nucleotides 45 to 2567) and BVH-11-2 (SEQ ID NO:13; ORF at nucleotides 114 to 2630).

The first amino acids of the BVH-3 and BVH-11 coding regions have the characteristics of leader sequences also known as signal peptides. The consensus signal peptidase cleavage site L-X-X-C of lipoprotein modification/processing sites was present in the sequences. Mature BVH-3, BVH-11 and BVH-11-2 proteins from S. pneumoniae SP64 have 1019, 821 and 819 amino acids, respectively. The regions of S. pneumoniae genes coding for mature BVH-3, termed BVH-3M, (nucleotides 1837 - 4896; SEQ. ID. NO: 11), BVH-11M (nucleotides 102-2567; SEQ. ID. NO: 12) and BVH-11-2M (nucleotides 171-2630; SEQ. ID. NO: 13), were amplified by PCR(DNA Thermal Cycler GeneAmp PCR system 2400 Perkin Elmer, San Jose, CA) from genomic DNA of 6 or 7 S. pneumoniae strains. Serogroup 6 S. pneumoniae SP64 and serogroup 9 SP63 clinical isolates were provided by the laboratoire de la santé publique du Québec, Sainte-Anne-de-Bellevue; serotype 4 strain JNR.7/87 was provided by Andrew Camilli, Tufts University School of Medicine, Boston; Rx1 strain, a nonencapsulated derivative of the type 2 strain D39 and the type 3 strains A66 and WU2 were provided by David E. Briles from University of Alabama, Birmingham and the type 3 clinical isolate P4241 was provided by the centre de recherche en infectiologie du

centre hospitalier de l'université Laval, Sainte-Foy. The sets of oligonucleotide primers OCRR479-OCRR480; HAMJ160-OCRR488 and HAMJ160-HAMJ186, that contained base extensions for the addition of restriction sites were used for the amplification of BVH-3, BVH-11 and BVH-11-2 gene, respectively, with the exception of BVH-11 gene from SP64 strain which was amplified using the set of primers consisting of HAMJ487 and OCRR488. Primer sequences are listed below (Table 2). PCR products were purified from agarose gel using a QIAquick gel extraction kit from QIAgen (Chatsworth, CA) and digested BglII-XbaI or BglII-HindIII (Pharmacia Canada Inc, Baie d'Urfé, Canada). Digestions were cleaned using a QIAquick PCR purification kit from QIAgen (Chatsworth, CA). The PCR products were ligated to the BglII-XbaI or BglII-HindIII pSL301 vector. The ligated products were transformed into E. coli strain DH5α [ϕ80 *lac*Z ΔM15 *end*A1 *rec*A1 *hsd*R17 ($r^-K^+$) *sup*E44 *thi*-1λ$^-$ *gyr*A96 *rel*A1 Δ(*lac*ZYA-*arg*F)U169] (Gibco BRL, Gaithersburg, MD) according to the method of Simanis (Hanahan, D. DNA Cloning, 1985, D.M. Glover (ed), pp. 109-135). Recombinant pSL301 plasmids (rpSL301) containing BVH-3, BVH-11 or BVH11-2 were purified using a QIAgen kit (Chatsworth, CA) and DNA inserts were sequenced (Taq Dye Deoxy Terminator Cycle Sequencing kit, ABI, Foster City, CA). The figures 11 and 12 depict the consensus sequence established from the BVH-3, and BVH-11 deduced amino acid sequences, respectively. Comparison of BVH-3 protein sequences revealed 99 to 100% identity of sequences for all strains with the exception that BVH-3 from serogroup 9 SP63 strain (SEQ. ID. NO: 15 and SEQ. ID. NO: 16) misses a stretch of 177 amino acids corresponding to residues 244 to 420 on BVH-3'protein sequence of S. pneumoniae SP64. Analysis of sequences of additional serogroup 9 strains revealed BVH-3 molecule having the same deletion in 3 out of 4 strains

thus suggesting that the 3 strains are members of a S. pneumoniae serogroup 9 clone.

Comparison of 13 BVH-11 nucleotide sequences obtained from 7 S. pneumoniae strains, revealed that the nucleotide sequences are very similar. Computer analysis (MacVector, Clustal W 1.4) using multiple alignment of the predicted BVH-11 protein sequences revealed that these sequences were 75% identical and 82 % homologous on a length of 834 amino acids. Pairwise alignment revealed 80 to 100% identity (Figure 13). The sequences showed great similarity in overall organization. Variability in the primary sequence of these proteins is almost restricted to the last 125 amino acids in the C-terminal portion of the proteins. This region constitutes a domain. Close examination of this domain revealed two groups of sequences. The first 9 sequences from the figure 13 belong to one group while the last 4 sequences belong to another group. A 39% identity value is obtained when the domain sequences of the 13 proteins are compared (MacVector, Clustal W 1.4). The identity value increased to more than 92% when sequences belonging to a same group are compared.

EXAMPLE 7

This example illustrates the homology of portions of BVH-3 and BVH-11 genes.

Molecular analysis with DNA probes derived from BVH-3 and BVH-11 genes indicated that BVH-3 and BVH-11 were related. In dot blot hybridization studies, DNA probe consisting of either, BVH-3 or BVH-11, gene sequence hybridized to both, BVH-3 and BVH-11 genes thus indicating that BVH-3 and BVH-11 genes shared homologous sequences. Comparison of sequences revealed that the ORFs and the proteins were 43

and 33% identical, respectively. Closer examination revealed that the region corresponding to amino acids 1 to 225 in BVH-3 and 1 to 228 in BVH-11 were 73 and 75% identical at the DNA and protein level, respectively. In contrast, the 3' regions corresponding to amino acids 226 to 1039 from BVH-3 and amino acids 229-840 from BVH-11 were only 34 and 22% identical at the DNA and protein level, respectively. Thus the 5' termini of BVH-3 and BVH-11 genes appear to contain highly conserved sequences while the remaining parts of the genes are highly divergent. These results suggest that BVH-3 and BVH-11 might share similar functions mediated by sequences present in the conserved region whereas BVH-3- and BVH-11-specific functions might be mediated by sequences in the divergent region.

EXAMPLE 8

This example describes the cloning of truncated BVH-3, BVH-11 and BVH-11-2 genes by polymerase chain reaction (PCR) and the expression of truncated BVH-3 and BVH-11 molecules.

Gene fragments were amplified by PCR using pairs of oligonucleotide engineered to amplify fragments spanning the BVH-3 (**SEQ ID NO: 1** and **SEQ ID NO: 11**), BVH-11 (**SEQ ID NO: 3** and **SEQ ID NO: 12**) or BVH-11-2 (**SEQ ID NO: 13**) gene from S. pneumoniae strain SP64. Each of the primers had a restriction endonuclease site at the 5' end, thereby allowing directional in-frame cloning of the amplified product into the digested plasmid vector (Tables 2 and 3). PCR-amplified products were digested with restriction endonucleases and ligated to either linearized plasmid pSL301 (see example 1), pCMV-GH (see example 2) or pET (Novagen, Madison, WI) expression vector digested likewise or digested with enzymes that produce compatible cohesive

ends. Recombinant pSL301 and recombinant pCMV-GH plasmids were digested with restriction enzymes for the in-frame cloning in pET expression vector. Clones were first stabilized in E. coli DH5α before introduction into E. coli BL21(λDE3) or AD494 (λDE3) for expression of truncated BVH-3 or BVH-11 molecules. Each of the resultant plasmid constructs was confirmed by nucleotide sequence analysis. The recombinant proteins were expressed as N-terminal fusions with the thioredoxin and His-tag or as C-terminal fusions with an His-tag. The expressed recombinant proteins were purified from supernatant fractions obtained from centrifugation of sonicated IPTG-induced E. coli cultures using a His-Bind metal chelation resin (QIAgen, Chatsworth, CA). The gene products generated are listed in the table 3. The gene products corresponding to the N-terminal region including the signal sequence are designated as Lipidated-proteins or lipoproteins (L-proteins). The gene products corresponding to the N-terminal region lacking the signal sequence are identified as protein without signal sequence (w/o ss).

Table 2. List of PCR oligonucleotide primers

| Primer | SEQ. ID. | Sequence 5' - 3' | Nucleotide position | Restric-tion sites |
|---|---|---|---|---|
| OCRR 479 | 17 | cagtagatctgtgcctatgcactaaac | SEQ ID 1 :61-78 | BglII |
| OCRR 480 | 18 | gatctctagactactgctattccttacgctatg | SEQ ID 11 :4909-4887 | XbaI |
| OCRR 497 | 19 | atcactcgagcattacctggataatcctgt | SEQ ID 1 :1525-1506 | XhoI |
| OCRR 498 | 20 | ctgctaagcttatgaaagatttagat | SEQ ID 1 :1534-1548 | HindIII |

| | | | | |
|---|---|---|---|---|
| OCRR 499 | 21 | gatactcgagctgctattccttac | SEQ ID 11 :4906-4893 | XhoI |
| HAMJ 172 | 22 | gaatctcgagttaagctgctgctaattc | SEQ ID 1 : 675-661 | XhoI |
| HAMJ 247 | 23 | gacgctcgagcgctatgaaatcagataaattc | SEQ ID 1 :3117-3096 | XhoI |
| HAMJ 248 | 24 | gacgctcgagggcattacctggataatcctgttcatg | SEQ ID 1 :1527-1501 | XhoI |
| HAMJ 249 | 25 | cagtagatctcttcatcatttattgaaaagagg | SEQ ID 11 : 1749-1771 | BglII |
| HAMJ 278 | 26 | ttatttcttccatatggacttgacagaagagcaaattaag | SEQ ID 1 :1414-1437 | NdeI |
| HAMJ 279 | 27 | cgccaagcttcgctatgaaatcagataaattc | SEQ ID 1 :3117-3096 | HindIII |
| HAMJ 280 | 28 | cgccaagcttttccacaatataagtcgattgatt | SEQ ID 1 :2400-2377 | HindIII |
| HAMJ 281 | 29 | ttatttcttccatatggaagtacctatcttggaaaaagaa | SEQ ID 1 :2398-2421 | NdeI |
| HAMJ 300 | 30 | ttatttcttccatatggtgcctatgcactaaaccagc | SEQ ID 1 :62-82 | NdeI |
| HAMJ 313 | 31 | ataagaatgcggccgcttccacaatataagtcgattgatt | SEQ ID 1 :2400-2377 | NotI |
| OCRR 487 | 32 | cagtagatctgtgcttatgaactaggtttgc | SEQ ID 3 :58-79 | BglII |
| OCRR 488 | 33 | gatcaagcttgctgctacctttacttactctc | SEQ ID 12 :2577-2556 | HindIII |
| HAMJ 171 | 34 | ctgagatatccgttatcgttcaaacc | SEQ ID 3 :1060-1075 | EcoRV |
| HAMJ 251 | 35 | ctgcaagcttttaaaggggaataatacg | SEQ ID 3 :1059-1045 | HindIII |
| HAMJ 264 | 36 | cagtagatctgcagaagccttcctatctg | SEQ ID 3 :682-700 | BglII |
| HAMJ 282 | 37 | tcgccaagcttcgttatcgttcaaaccattggg | SEQ ID 3 :1060-1081 | HindIII |
| HAMJ 283 | 38 | ataagaatgcggccgccttactctcctttaataaagccaat agtt | SEQ ID 3 :2520-2492 | NdeI |
| HAMJ 284 | 39 | catgccatggacattgatagtctcttgaaacagc | SEQ ID 3 :856-880 | NcoI |
| HAMJ 285 | 40 | cgccaagcttcttactctcctttaataaagccaatag | SEQ ID 3 :2520-2494 | HindIII |
| HAMJ 286 | 41 | cgacaagcttaacatggtcgctagcgttacc | SEQ ID 3 :2139-2119 | HindIII |
| HAMJ 287 | 42 | cataccatgggcctttatgaggcacctaag | SEQ ID 3 :2014-2034 | NcoI |
| HAMJ 288 | 43 | cgacaagcttaagtaaatcttcagcctctctcag | SEQ ID 3 :2376-2353 | HindIII |

| | | | | |
|---|---|---|---|---|
| HAMJ 289 | 44 | gataccatggctagcgaccatgttcaaagaa | SEQ ID 3 :2125-2146 | NcoI |
| HAMJ 290 | 45 | cgccaagcttatcatccactaacttgactttatcac | SEQ ID 3 :1533-1508 | HindIII |
| HAMJ 291 | 46 | cataccatggatattcttgccttcttagctccg | SEQ ID 3 :1531-1554 | NcoI |
| HAMJ 301 | 47 | catgccatggtgcttatgaactaggtttgc | SEQ ID 3 :59-79 | NcoI |
| HAMJ 302 | 48 | cgccaagctttagcgttaccaaaaccattatc | SEQ ID 3 :2128-2107 | HindIII |
| HAMJ 160 | 49 | gtattagatctgttcctatgaacttggtcgtcacca | SEQ ID 13 : 172-196 | BglII |
| HAMJ 186 | 50 | cgcctctagactactgtataggagccgg | SEQ ID 13: 2460-2443 | XbaI |
| HAMJ 292 | 51 | catgccatggaaaacatttcaagccttttacgtg | SEQ ID 11: 754-778 | NcoI |
| HAMJ 293 | 52 | cgacaagcttctgtataggagccggttgactttc | SEQ ID 11 : 2457-2434 | HindIII |
| HAMJ 294 | 53 | catgccatggttcgtaaaaataaggcagaccaag | SEQ ID 11 : 2038-2062 | NcoI |
| HAMJ 297 | 54 | catgccatggaagcctattggaatgggaag | SEQ ID 11 : 622-642 | NcoI |

Table 3. Lists of truncated BVH-3 and BVH-11 gene products generated from S. pneumoniae SP64

| PCR-primer sets | Protein designation | Identification (encoded amino acids) | SEQ. ID.NO. | Cloning vector |
|---|---|---|---|---|
| OCRR479-OCRR480 | BVH-3M | BVH-3 w/o ss (21-1039) | 55 | pSL301 |
| OCRR479-OCRR497 | BVH-3AD | BVH-3 N'end w/o ss (21-509) | 56 | pSL301 |
| HAMJ248-HAMJ249 | L-BVH-3AD | BVH-3 N'end (1-509) | 57 | pET-21(+) |
| OCRR498-OCRR499 | BVH-3B | BVH-3 C'end (512-1039) | 10 | pSL301 |
| OCRR479-HAMJ172 | BVH-3C | BVH-3 N'end w/o ss (21-225) | 59 | pET-32 c(+) |
| OCRR487-OCRR488 | BVH-11M | BVH-11 w/o ss (20-840) | 60 | pCMV-GH |
| HAMJ251-OCRR487 | BVH-11A | BVH-11 N'end w/o ss (20-353) | 61 | pET-32 c (+) |
| HAMJ171-OCRR488 | BVH-11B | BVH-11 C'end (354-840) | 62 | pET-32 a(+) |
| HAMJ264-OCRR488 | BVH-11C | BVH-11 C'end (228-840) | 63 | pET-32 a(+) |
| HAMJ278-HAMJ279 | NEW1 | BVH-3 C'end (472-1039) | 64 | pET-21b(+) |
| HAMJ278-HAMJ280 | NEW2 | BVH-3 C'end (472-800) | 65 | pET-21b(+) |
| HAMJ281-HAMJ279 | NEW3 | BVH-3 C'end (800-1039) | 66 | pET-21b(+) |
| HAMJ284-HAMJ285 | NEW4 | BVH-11 C'end (286-840) | 67 | pET-21d(+) |
| HAMJ284-HAMJ286 | NEW5 | BVH-11 internal (286-713) | 68 | pET-21d(+) |
| HAMJ287-HAMJ288 | NEW6 | BVH-11 internal (672-792) | 69 | pET-21d(+) |
| HAMJ285-HAMJ289 | NEW7 | BVH-11 internal (709-840) | 70 | pET-21d(+) |
| HAMJ284-HAMJ290 | NEW8 | BVH-11 internal (286-511) | 71 | pET-21d(+) |

| | | | | |
|---|---|---|---|---|
| HAMJ286-HAMJ291 | NEW9 | BVH-11 internal (511-713) | 72 | pET-21d(+) |
| HAMJ160-HAMJ186 | BVH-11-2M | BVH-11-2 w/o ss (20-838) | 73 | pSL301 |
| HAMJ292-HAMJ293 | NEW10 | BVH-11-2 C'end (271-838) | 74 | pET-21d(+) |
| HAMJ293-HAMJ294 | NEW11 | BVH-11-2 C'end (699-838) | 75 | pET-21d(+) |
| HAMJ282-HAMJ283 | BVH-11B | BVH-11 C'end (354-840) | 62 | pET-21b(+) |
| HAMJ286-HAMJ297 | NEW14 | BVH-11-2 internal (227-699) | 77 | pET-21d(+) |
| HAMJ300-HAMJ313 | NEW15 | BVH-3 N'end w/o ss (21-800) | 78 | pET-21b(+) |
| HAMJ301-HAMJ302 | NEW16 | BVH-11 N'end w/o ss (20-709) | 79 | pET-21d(+) |

EXAMPLE 9

This example describes the isolation of monoclonal antibodies (Mabs) and the use of Mabs to characterize BVH-3, BVH-11 and BVH-11-2 protein epitopes.

Female BALB/c mice (Charles River) were immunized subcutaneously with BVH-3, BVH-11 or BVH-11-2 gene products from S. pneumoniae strain SP64 in presence of 15 μg of QuilA adjuvant (Cedarlane Laboratories Ltd, Hornby, Canada). One set of mice (fusion experiment 1) were immunized on day 1 and 14 with 25 μg of affinity purified thioredoxin-His•Tag-BVH-3M fusion protein. A second group of mice (fusion experiment 2) were immunized three times at three-week intervals with 25 μg of affinity purified thioredoxin-His•Tag-BVH-11M. A third group of mice (fusion experiment 3) were immunized on day 1 and day 15 with 25 μg of affinity purified thioredoxin-His•Tag-BVH-11-2M fusion protein. A fourth group of mice (fusion experiment 4) were immunized on day 1 with 25 μg of affinity purified thioredoxin-His•BVH-11B fusion protein and boosted by intravenous injection on day 16 and on day 37 with recombinant BVH-11B in PBS. Three to four days before fusion, mice were injected intravenously with 25 μg of the respective antigen suspended in PBS alone. Hybridomas were produced by fusion of spleen cells with nonsecreting SP2/0 myeloma cells as previously described by J. Hamel et al. [J. Med. Microbiol., 23, pp163-170 (1987)]. Culture supernatants of hybridomas were initially screened by enzyme-linked-immunoassay according to the procedure described by Hamel et al. (Supra) using plates coated with preparations of purified recombinant proteins or suspensions of heat-killed S. pneumoniae cells. Positive hybridomas selected on the basis of ELISA reactivity with a

variety of antigens were then cloned by limiting dilutions, expanded and frozen.

Hybridomas were tested by ELISA or Western immunoblotting against BVH-3 and BVH-11 gene products in order to characterize the epitopes recognized by the Mabs. BVH-3 and BVH-11 shared common epitopes with 6 Mabs (H3-1-F9, H3-1-D4, H3-1-H12, H11-1-E7, H11-1-H10 and H11-1.1-G11) showing reactivities with both proteins (Table 4). BVH-11 and BVH-11-2 molecules from S. pneumoniae SP64 shared common epitopes not present on BVH-3 with Mabs (3A1, 13C11, 10H10, 1D8, 10G9, 10A2, 3E8, 10D7, 2H7 and 6H7) reactive with both, BVH-11 and BVH-11-2, recombinant proteins (Table 5).

Table 4. Reactivity of BVH-3-immunoreactive Mabs with a panel of BVH-3 and BVH-11 gene products

| MAbs | a. Immunoreactivity with | | | | | | |
|---|---|---|---|---|---|---|---|
| | BVH-3M 21-1039 | BVH-3A 21-509 | BVH-3B 512-1039 | BVH-3C 21-225 | NEW2 472-800 | NEW3 800-1039 | BVH-11M 20-840 |
| H3-1-F9 | + | + | - | + | - | - | + |
| H3-1-D4 | + | + | - | + | - | - | + |
| H3-1-H12 | + | + | - | + | - | - | + |
| H3-2-G2 | + | + | - | - | - | - | - |
| H3-3-A1 | + | + | - | - | - | - | - |
| H3-4-D3 | + | - | + | - | - | + | - |
| H11-1-E7 | + | + | - | + | - | - | + |
| H11-1-H10 | + | + | - | + | - | - | + |
| H11-1.1-G11 | + | + | - | + | + | - | + |

Table 5. Reactivity of Mabs raised against BVH-11-2 protein from S. pneumoniae strain SP64 with a panel of BVH-11 gene products

| Mabs[a] | b. Immunoreactivity with | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | c. BVH-11 products | | | | d. BVH-11-2 products | | | |
| | BVH-11M 20-840 | NEW8 286-511 | NEW9 511-713 | BVH-11B 354-840 | BVH-11-2 20-838 | NEW10 271-838 | NEW11 699-838 | NEW14 227-699 |
| 3A1 | + | + | - | + | + | + | - | + |
| 13C1 | + | + | + | + | + | + | - | + |
| 10H10 | + | + | + | + | + | + | - | + |
| 1D8 | + | + | - | + | + | + | - | + |
| 10G9 | + | - | - | + | + | + | - | + |
| 10A2 | + | - | - | + | + | + | - | + |
| 3E8 | + | - | - | + | + | + | - | + |
| 10D7 | + | - | - | + | + | + | - | + |
| 2H7 | + | - | - | - | + | - | - | - |
| 6H7 | + | - | - | - | + | - | - | - |
| 3A4 | - | - | - | - | + | + | + | - |
| 14H6 | - | - | - | - | + | + | + | - |
| 7G2 | - | - | - | - | + | + | - | + |
| 13H10 | - | - | - | - | + | - | - | + |
| 7E8 | - | - | - | - | + | - | - | - |
| 7H6 | - | - | - | - | + | - | - | - |

[a] Mabs listed in this table were not reactive with recombinant BVH-3 molecule

The results obtained from the immunoreactivity studies of the Mabs (Table 4 and Table 5) are in agreement with the protein sequences derived from the respective gene sequences. Indeed the Mabs cross-reactive with BVH-3 and BVH-11 molecules recognized BVH-3C protein corresponding to the conserved region, and BVH-11 and BVH-11-2 specific Mabs were reactive with epitopes located on variable parts of these molecules. BVH-3 and BVH-11, and BVH-11 and BVH-11-2 can be distinguished by their reactivity with Mabs.

EXAMPLE 10

This example illustrates the simultaneous expression of BVH-3 and BVH-11 gene products by S. pneumoniae.

A standard Western blot technique was used to investigate whether BVH-3 and BVH-11 genes were expressed in S. pneumoniae. S. pneumoniae strain SP64 and SP63 were grown overnight at 37°C in 5% $CO_2$ on chocolate agar plates, bacteria were suspended in PBS and heat-killed at 56°C for 20 min. For the preparation of antigens, suspensions of S. pneumoniae were treated with sample buffer containing SDS and 2-mercaptoethanol for 5 min at 100°C. Pneumococcal protein antigens were resolved by SDS-PAGE electrophoresis according to the method of Laemmli [Nature, 227, pp. 680-685 (1970)]. After SDS-PAGE, the proteins were transferred electrophoretically from the gel to nitrocellulose paper by the method of Towbin [Proc. Natl. Acad. Sci. USA, 76, pp. 4350-4354 (1979)] and probed with mouse antiserum or monoclonal antibodies. The detection of antigens reactive with the antibodies was performed by indirect enzyme-immunoassay using conjugated-anti-mouse immunoglobulins and a colour substrate. When antiserum raised to recombinant BVH-3 was tested against S. pneumoniae SP64 antigens, two reactive bands having apparent molecular masses of 127 kDa and 99 kDa were detected. Bands having the same apparent molecular masses were also detected when Mabs H3-1-F9, H3-1-D4, H3-1-H12, H11-1-E7, H11-1-H10 and H11-1.1-G11 were used individually as immunological probes. In contrast, Mabs specific for the BVH-3 molecule detected the 127 kDa band only and Mabs specific for BVH-11 detected the 99 kDa band only thus confirming the identity of the 127 and 99 kDa bands as BVH-3 and BVH-11, respectively. These studies provide evidence that BVH-3 and BVH-11 proteins are simultaneously present on S. pneumoniae. Moreover, the results are consistent with our previous observations that BVH-3 and BVH-11 possess epitopes that are common to both proteins and epitopes that are exclusive to either protein.

In S. pneumoniae SP64, mature BVH-3, BVH-11 and BVH-11-2 are proteins of 1019, 821 and 819 amino acids with predicted molecular mass of 112.5 kDa, 92.4 kDa, and 91.7 kDa, respectively. Although there is a discrepancy between the molecular mass predicted from the sequence and the molecular mass calculated on SDS-PAGE, BVH-3 can be distinguished from BVH-11 by its higher molecular mass. Moreover, BVH-3 molecules from S. pneumoniae strain SP63 have an apparent molecular mass of 112 kDa in SDS-PAGE compared to 127 kDa for BVH-3 of SP64 strain. This data is consistent with the deletion of a stretch of 177 amino acid residues in BVH-3 of S. pneumoniae strain SP63.

EXAMPLE 11

This example describes the protection conferred in experimental infection of mice vaccinated with recombinant BVH-3 or BVH-11 gene products.

Groups of 7 or 8 female BALB/c mice (Charles River) were immunized subcutaneously three times at three-week intervals with either affinity purified thioredoxin-His•Tag-BVH-3M fusion protein, affinity purified thioredoxin-His•Tag-BVH-11M fusion protein or, as control, with QuilA adjuvant alone in PBS. Twelve to 14 days following the third immunization, the mice were challenged intravenously with S. pneumoniae WU2 strain or intranasally with P4241 strain. Samples of the S. pneumoniae challenge inoculum were plated on chocolate agar plates to determine the CFU and to verify the challenge dose. The challenge dose was approximately $10^6$ CFU. Deaths were recorded for a period of 14 days and on day 14 post-challenge, the surviving mice were sacrificed and blood samples tested for

the presence of S. pneumoniae organisms. The survival data are shown in Tables 6 and 7.

Table 6. Protection mediated by recombinant BVH-3M and BVH-11M proteins in experimental infection with virulent S. pneumoniae WU2

| Experiment | Immunogen | Alive : dead[a] | Median days alive |
|---|---|---|---|
| 1 | BVH-3M | 8 : 0 | >14 |
| | none | 0 : 8 | 1 |
| 2 | BVH-11M | 8 : 0 | >14 |
| | none | 0 : 8 | 1 |

[a] The number of mice alive : the number of mice dead on day 14 post-challenge.

Table 7. Protection mediated by recombinant BVH-3M and BVH-11M proteins in experimental pneumonia with virulent S. pneumoniae P4241

| Experiment | Immunogen | Alive : dead[a] | Median day alive |
|---|---|---|---|
| 1 | BVH-3M | 6 : 1 | >14 |
| | none | 1 : 7 | 4.5 |
| 2 | BVH-3M | 8 : 0 | >14 |
| | BVH-11M | 8 : 0 | >14 |
| | none | 0 : 8 | 4 |

[a] The number of mice alive : the number of mice dead on day 14 post-challenge.

All mice immunized with recombinant BVH-3M or BVH-11M protein survived to infection with WU2 while none of the control mice given adjuvant alone survived. All except one mice immunized with recombinant BVH-3M or BVH-11M protein survived to infection with P4241 while only one control mice given adjuvant alone survived. All hemocultures from

surviving mice were negative at day 14 post-challenge. These results clearly indicate that both, BVH-3M and BVH-11M, elicit protective anti-pneumococcal immune responses in mice. The fact that these proteins are highly conserved among S. pneumoniae isolates emphasize the potential of BVH-3 and BVH-11 as universal vaccine candidates. Indeed, the BVH-3 and BVH-11 proteins from serogroup 6 S. pneumoniae strain SP64 elicited protection against pneumococcal infections with strains of different capsular serotypes.

Ideally, a vaccine that could protect against pneumococcal disease, could protect against meningitis, otitis media, bacteremia and pneumonia. BVH-3 and BVH-11 were protective against lethal systemic- and pneumonia-infection models thus suggesting that, in humans, BVH-3- and BVH11-protein-based vaccines could reduce the incidence of a wide spectrum of disease caused by virtually all S. pneumoniae independently of the capsular serotype.

Data from Tables 6 and 7 clearly demonstrate that BVH-3 and BVH-11 were, both, protection-eliciting molecules of S. pneumoniae. It was not known, however, whether protection can be mediated by specific sequences that were not shared on BVH-3 and BVH-11 molecules. Groups of female BALB/c mice (Charles River) were immunized subcutaneously three times at three-week intervals with either affinity purified thioredoxin-His•Tag- BVH-3AD, -BVH-3B or -BVH-3C fusion protein in presence of 15 μg of QuilA adjuvant (Cedarlane Laboratories Ltd, Hornby, Canada). Control mice were immunized with QuilA adjuvant alone in PBS or affinity purified thioredoxin-His•Tag or thioredoxin-His•Tag-fusion protein (His-Thio) in presence of QuilA.

To determine the protective ability of a set of truncated proteins, termed NEW4, NEW5, NEW6, NEW7, NEW8, NEW9, NEW10,

NEW11, NEW14 and BVH-11B, groups of female BALB/c mice (Charles River) were immunized subcutaneously two times at three-week intervals with 25 μg of either affinity purified His•Tag-fusion protein in presence of 15 μg of QuilA adjuvant. Ten to 14 days following the last immunization, the mice were challenged with virulent S. pneumoniae. Our results indicate that, BVH-3B, a truncated BVH-3 molecule consisting of amino acids 512-1039, elicited protection against the mouse-virulent strains WU2 and P4241. Similarly, BVH-11B, NEW4 and NEW5 molecules, three truncated BVH-11 molecules consisting of amino acids 354-840, amino acids 286-840 and amino acids 286-713, respectively, elicited protection against experiment intravenous challenge with WU2 and intranasal challenge with P4241. Moreover, vaccination with NEW10 and NEW14, consisting of amino acids 272-838 and amino acids 227-699 from BVH-11-2 molecule also resulted in protection against death with the pneumococcal strains. These results indicate that the region comprising 428 amino acids extending from amino acids 286-713 and amino acids 272-699 on S. pneumoniae SP64 BVH-11 and BVH-11-2 protein sequences, respectively, contains protective epitopes. This region is highly conserved with a global 91% identity and 94% homology among thirteen BVH-11 protein sequences.

Table 8. Evaluation of protection elicited by vaccination of mice with BVH-3 and BVH-11 gene products

| | | Challenge with WU2 | | Challenge with P4241 | |
|---|---|---|---|---|---|
| Experiment | Immunogen | Alive : dead[a] | Median day alive | Alive : dead | Median day alive |
| 1[b] | None | 0 : 8 | 1.5 | 1 : 7 | 4.5 |
| | NEW4 | 8 : 0 | >14 | 8 : 0 | >14 |
| | NEW5 | 8 : 0 | >14 | 8 : 0 | >14 |
| | NEW7 | 0 : 8 | 2 | 0 : 8 | 5 |
| | BVH-11M | 8 : 0 | >14 | 8 : 0 | >14 |
| 2[b] | None | 0 : 8 | 1 | 0 : 8 | 4 |
| | NEW5 | 8 : 0 | >14 | 8 : 0 | >14 |
| | NEW8 | 0 : 8 | 1.5 | 0 : 8 | 5.5 |
| | NEW9 | 3 : 5 | 3.5 | 2 : 6 | 7 |
| | BVH-11M | 8 : 0 | >14 | 8 : 0 | >14 |
| 3[b] | None | 0 : 8 | 1 | 0 : 8 | 4 |
| | NEW6 | 0 : 8 | 1 | 4 : 4 | 10.5[c] |
| | NEW10 | 8 : 0 | >14 | 8 : 0 | >14 |
| | NEW11 | 0 : 8 | 1.5 | 1 : 7 | 6 |
| | BVH-11M | 8 : 0 | >14 | 8 : 0 | >14 |
| 4[b] | None | 0 : 8 | 2 | 0 : 8 | 4 |
| | BVH-11B | 7 : 1 | >14 | 8 : 0 | >14 |
| | NEW14 | 8 : 0 | >14 | 8 : 0 | >14 |
| 5 | His-Thio | 0 : 8 | 2 | | |
| | BVH-3AD | 1 : 7 | 2.5 | | |
| | BVH-3B | 5 : 3 | >14 | | |
| 6 | His-Thio | 0 : 8 | 1 | | |
| | BVH-3C | 0 : 8 | 1 | | |

[a] The number of mice alive : the number of mice dead on day 14 post-challenge.

[b] The WU2 challenge dose was $10^5$ CFU.

[c] Mice living longer than 14 days were assigned a survival time of 14 days for the determination of median values.

EXAMPLE 12

This example described the cloning and expression of a chimeric gene encoding for a chimeric polypeptide corresponding to the carboxy-terminal region of BVH-3 in fusion at the C' end to the carboxy-terminal region of BVH-11 and the additive protection observed after vaccination with a chimeric polypeptide.

It is clear from the studies described above that BVH-3 and BVH-11 are serologically distinct molecules simultaneously present on S. pneumoniae. The results of immunological studies of mice indicate that both proteins are good vaccine candidates. These proteins have the potential to provide protection against all pneumococci, regardless of serotype. Even though the two proteins share epitopes and sequences, they have different characteristics and may serve different biological functions. Thus, immunization against the two proteins may provide a higher level of protection than that imparted by each individually. To examine this, several avenues where full-length or truncated BVH-3 and BVH-11 are administered in combination or in conjugation can be explored. Here we describe the genetic engineering of a BVH-3-BVH-11 fusion gene and protein, termed NEW12 (**SEQ ID NO:76 and SEQ ID NO:58**, respectively), and the potential use of NEW12 protein as a vaccine.

BVH-3 and BVH-11 gene fragments corresponding to the 3'end of the genes were amplified by PCR using pairs of oligonucleotides engineered to amplify fragments spanning nucleotides 1414 to 3117(**SEQ ID NO: 1**) and nucleotides 1060 to 2520 (**SEQ ID NO: 3**) from S. pneumoniae strain SP64 BVH-3 and BVH-11 genes, respectively. The primers used, HAMJ278 and HAMJ279; HAMJ282 and HAMJ283 had a restriction

endonuclease site at the 5' end, thereby allowing directional in-frame cloning of the amplified product into the digested pET21b(+) plasmid vector (Table 2). PCR-amplified products were digested with restriction endonucleases and ligated to linearized plasmid pET21b(+) vector digested likewise. The resultant plasmid constructs were confirmed by nucleotide sequence analysis. The recombinant pET21b(+) plasmid containing the NdeI-HindIII BVH-3 PCR product was linearized by digestion with the restriction enzymes HindIII and NotI for the in-frame cloning of the HindIII-NotI DNA fragment obtained from the recombinant pET21(+) vector containing the BVH-11 gene fragment. Clones were first stabilized in E. coli DH5α before introduction into E. coli BL21(λDE3) for expression of a chimeric pneumococcal protein molecule. The recombinant chimeric polypeptide, termed NEW 12, was expressed as C-terminal fusion with an His-tag. The expressed recombinant NEW 12 protein was purified from supernatant fractions obtained from centrifugation of sonicated IPTG-induced E. coli cultures using a His-Bind metal chelation resin (QIAgen, Chatsworth, CA).

According to the same procedure described above, it is possible to construct other chimeric polypeptides, as a result of a simultaneous expression of New 1 and New 4, New 1 and New 5, New 1 and New 10, or New 1 and New 14. The construction can be with New 1 upstream or downstream of New 4, New 5, New 10, BVH-11B or New 14. It is also possible to construct other chimeric polypeptides as a result of a simultaneous expression of more than two fragments of either genes of BVH-3, BVH-11 or BVH-11-2.

Groups of 8 female BALB/c mice (Charles River) were immunized subcutaneously two times at three-week intervals with 25 μg of either affinity purified His•Tag-fusion NEW1,

BVH-11B or NEW12 protein in presence of 15 μg of QuilA adjuvant. Ten to 14 days following the last immunization, the mice were challenged with virulent S. pneumoniae. As demonstrated before, NEW1 and BVH-11B molecules comprising amino acids 472 to 1039 from BVH-3 protein and amino acids 354-840 from BVH-11 protein, respectively, correspond to portions of the proteins capable of eliciting a protective immune response. To determine if a chimeric polypeptide would significantly improve the protection compared with those seen for the individual counterparts, the challenge dose was adjusted in a manner that protection was not expected with NEW1 and BVH-11B molecules. Interestingly, the chimeric NEW12 protein, elicited protection against the mouse-virulent strains WU2 and P4241. Seven out of 8 mice immunized with NEW12 were still alive 10 days after the challenge while 28 out of 32 mice immunized with NEW1, BVH-11B, BVH-3M or adjuvant alone were dead by five days post-challenge. Thus, vaccination of mice with NEW12 provided the highest degree of protection against WU2 challenge. These results indicate that immunization with a chimeric polypeptide and possibly a combination of BVH-3 and BVH-11 gene products can provide additional protection to that obtained by administration of BVH-3 or BVH-11 antigens alone.

Table 9. Evaluation of protection elicited by vaccination of mice with the chimeric NEW12 molecule

| | Challenge with WU2 | | Challenge with P4241 | |
|---|---|---|---|---|
| Immunogen | Alive : dead[a] | Median day alive | Alive : dead | Median day alive |
| None | 0 : 8 | 1 | 0 : 8 | 5 |
| NEW1' | 2 : 6 | 2 | 1 : 7 | 8 |
| BVH-11B | 1 : 7 | 3.5 | 8 : 0 | >14 |
| NEW12 | 6 : 2 | >14 | 7 : 1 | >14 |
| BVH-3M | 1 : 7 | 3 | 8 : 1 | >14 |

EXAMPLE 13

This example illustrates the identification of additional BVH-3 and BVH-11 related sequences in Streptococcus species other than S. pneumoniae.

It was previously shown that BVH-3, BVH-11 and BVH-11-2 are a family of related proteins sharing common sequences. Homology searches were performed with the nucleotide sequence from the conserved region of these genes and compared with GenBank and EMBL sequences using FASTA. The most significant homology was observed with a 2.469-kb gene coding for a calculated 92-kDa protein (**SEQ ID NO: 81**) of unknown function in S. agalactiae also called group B streptococcus or GBS. The gene was designated BVH-71. A protein demonstrating 99.2% identity and 99.5% similarity with that of GBS was also identified in S. pyogenes also called group A streptococcus or GAS (**SEQ ID NO: 83**). The 5' region of the BVH-71 sequences (**SEQ ID NO: 80** and **SEQ ID NO: 82**), spanning nucleotides 1 to 717, demonstrated 58 and 60% identity with the conserved regions of BVH-3 (nucleotides 1 to 675) and BVH-11 (nucleotides 1 to 684) genes respectively. The first 239 amino acids of the translated sequences of the GBS and GAS BVH-71 open reading frames are 51 and 54% identical to the first 225 and 228 amino acids of BVH-3 and BVH-11, respectively. In addition to structural similarities, streptococcal BVH-3, BVH-11 and BVH-71 proteins also share antigenic epitopes. A 97-kDa band was revealed on Western blots of GAS or GBS whole cells, using Mab H11-1.1-G11 reactive with the BVH-3 and BVH-11 conserved regions. Similarly, GAS and GBS

recombinant BVH-71 proteins were detected in Western immunoblot analysis.

These results indicate that BVH-71, BVH-3 and BVH-11 proteins might share similar functions. Our results also suggest that BVH-71 proteins can be used as protein vaccine components of anti-streptococcus. In a further embodiment BVH-71 proteins can be used as protein vaccine components of anti-GAS or anti-GBS vaccines.

What is claimed is:

1. An isolated polynucleotide encoding a polypeptide having at least 70% identity to a second polypeptide having a sequence chosen from: **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

2. A polynucleotide according to claim 1, wherein said polynucleotide encodes a polypeptide having at least 95% identity to the second polypeptide.

3. An isolated polynucleotide encoding a polypeptide capable of generating antibodies having binding specificity for a polypeptide having a sequence chosen from: **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

4. An isolated polynucleotide that is complementary to the polynucleotide of claim 1.

5. An isolated polynucleotide that is complementary to the polynucleotide of claim 3.

6. The polynucleotide of claim 1, wherein said polynucleotide is DNA.

7. The polynucleotide of claim 3, wherein said polynucleotide is DNA.

8. The polynucleotide of claim 1, wherein said polynucleotide is RNA.

9. The polynucleotide of claim 3, wherein said polynucleotide is RNA.

10. A vector comprising the polynucleotide of claim 1, wherein said DNA is operably linked to an expression control region.

11. A vector comprising the polynucleotide of claim 3, wherein said DNA is operably linked to an expression control region.

12. A host cell transfected with the vector of claim 10.

13. A host cell transfected with the vector of claim 11.

14. A process for producing a polypeptide comprising culturing a host cell according to claim 12 under conditions suitable for expression of said polypeptide.

15. A process for producing a polypeptide comprising culturing a host cell according to claim 13 under condition suitable for expression of said polypeptide.

16. An isolated polypeptide having at least 70% identity to a second polypeptide having an amino acid sequence chosen from: **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

17. An isolated polypeptide capable of generating antibodies having binding specificity for a second polypeptide having a sequence chosen from: **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

18. An isolated polypeptide having an amino acid sequence chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

19. An isolated polypeptide according to claim 18, wherein the N-terminal Met residue is deleted.

20. An isolated polypeptide according to claim 18, wherein the secretory amino acid sequence is deleted.

21. A chimeric polypeptide comprising two or more polypeptides chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof; provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

22. A chimeric polypeptide comprising two or more polypeptides chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof; provided that the polypeptides or fragments, analogs or derivatives thereof are linked as to form a chimeric polypeptide.

23. A chimeric polypeptide of formula (I):
**A-(B)$_m$-(C)$_n$-D** (I)
Wherein;
**m** is 0 or 1,
**n** is 0 or 1,
**A** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof;
**B** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof;
**C** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55 to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof; and
**D** is chosen from **SEQ ID NOs: 2, 4, 6, 8, 10, 14, 16, 55**

**to 75, 77 to 79, 81, 83** or fragments, analogs or derivatives thereof.

24. A chimeric polypeptide of formula (I):
**A-(B)$_m$-(C)$_n$-D** (I)
Wherein;
**m** is 0 or 1,
**n** is 0 or 1,
**A** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof;
**B** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77**, or fragments, analogs or derivatives thereof;
**C** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof; and
**D** is chosen from **SEQ ID NOs :10, 58, 60, 62, 64, 67, 68, 69, 72, 74, 77** or fragments, analogs or derivatives thereof.

25. A vaccine composition comprising a polypeptide according to any one of claims 16 to 24 and a pharmaceutically acceptable carrier, diluent or adjuvant.

26. A method for therapeutic or prophylactic treatment of meningitis, otitis media, bacteremia or pneumonia infection in an individual susceptible to meningitis, otitis media, bacteremia or pneumonia infection comprising administering to said individual a therapeutic or prophylactic amount of a composition according to claim 25.

27. A method for therapeutic or prophylactic treatment of streptococcal bacterial infection in an individual

susceptible to streptococcal infection comprising administering to said individual a therapeutic or prophylactic amount of a composition according to claim 25.

28. A method according to claim 26, wherein said individual is a mammal.

29. A method according to claim 27, wherein said individual is a human.

30. A method according to claim 22, wherein said bacterial infection is S.pneumoniae, group A *streptococcus (pyogenes)*, group B *streptococcus* (GBS or *agalactiae*), *dysgalactiae*, *uberis*, *nocardia* or *Staphylococcus aureus*.

31. A method according to claim 26, wherein said bacterial infection is S.pneumoniae.

32. Use of a vaccine composition according to claim 25 for the prophylactic or therapeutic treatment of Streptococcal infection in an animal susceptible to or infected with streptococcal infection comprising administering to said animal a prophylactic or therapeutic amount of the composition.

```
ATGAAATTTA GTAAAAAATA TATAGCAGCT GGATCAGCTG TTATCGTATC CTTGAGTCTA      60
TGTGCCTATG CACTAAACCA GCATCGTTCG CAGGAAAATA AGGACAATAA TCGTGTCTCT     120
TATGTGGATG GCAGCCAGTC AAGTCAGAAA AGTGAAAACT TGACACCAGA CCAGGTTAGC     180
CAGAAAGAAG GAATTCAGGC TGAGCAAATT GTAATCAAAA TTACAGATCA GGGCTATGTA     240
ACGTCACACG GTGACCACTA TCATTACTAT AATGGGAAAG TTCCTTATGA TGCCCTCTTT     300
AGTGAAGAAC TCTTGATGAA GGATCCAAAC TATCAACTTA AAGACGCTGA TATTGTCAAT     360
GAAGTCAAGG GTGGTTATAT CATCAAGGTC GATGGAAAAT ATTATGTCTA CCTGAAAGAT     420
GCAGCTCATG CTGATAATGT TCGAACTAAA GATGAAATCA ATCGTCAAAA ACAAGAACAT     480
GTCAAAGATA ATGAGAAGGT TAACTCTAAT GTTGCTGTAG CAAGGTCTCA GGGACGATAT     540
ACGACAAATG ATGGTTATGT CTTTAATCCA GCTGATATTA TCGAAGATAC GGGTAATGCT     600
TATATCGTTC CTCATGGAGG TCACTATCAC TACATTCCCA AAAGCGATTT ATCTGCTAGT     660
GAATTAGCAG CAGCTAAAGC ACATCTGGCT GGAAAAAATA TGCAACCGAG TCAGTTAAGC     720
TATTCTTCAA CAGCTAGTGA CAATAACACG CAATCTGTAG CAAAAGGATC AACTAGCAAG     780
CCAGCAAATA AATCTGAAAA TCTCCAGAGT CTTTTGAAGG AACTCTATGA TTCACCTAGC     840
GCCCAACGTT ACAGTGAATC AGATGGCCTG GTCTTTGACC CTGCTAAGAT TATCAGTCGT     900
ACACCAAATG GAGTTGCGAT TCCGCATGGC GACCATTACC ACTTTATTCC TTACAGCAAG     960
CTTTCTGCTT TAGAAGAAAA GATTGCCAGA ATGGTGCCTA TCAGTGGAAC TGGTTCTACA    1020
GTTTCTACAA ATGCAAAACC TAATGAAGTA GTGTCTAGTC TAGGCAGTCT TTCAAGCAAT    1080
CCTTCTTCTT TAACGACAAG TAAGGAGCTC TCTTCAGCAT CTGATGGTTA TATTTTTAAT    1140
CCAAAAGATA TCGTTGAAGA AACGGCTACA GCTTATATTG TAAGACATGG TGATCATTTC    1200
CATTACATTC CAAAATCAAA TCAAATTGGG CAACCGACTC TTCCAAACAA TAGTCTAGCA    1260
ACACCTTCTC CATCTCTTCC AATCAATCCA GGAACTTCAC ATGAGAAACA TGAAGAAGAT    1320
GGATACGGAT TTGATGCTAA TCGTATTATC GCTGAAGATG AATCAGGTTT TGTCATGAGT    1380
CACGGAGACC ACAATCATTA TTTCTTCAAG AAGGACTTGA CAGAAGAGCA AATTAAGGCT    1440
GCGCAAAAAC ATTTAGAGGA AGTTAAAACT AGTCATAATG GATTAGATTC TTTGTCATCT    1500
CATGAACAGG ATTATCCAGG TAATGCCAAA GAAATGAAAG ATTTAGATAA AAAAATCGAA    1560
GAAAAAATTG CTGGCATTAT GAAACAATAT GGTGTCAAAC GTGAAAGTAT TGTCGTGAAT    1620
AAAGAAAAAA ATGCGATTAT TTATCCGCAT GGAGATCACC ATCATGCAGA TCCGATTGAT    1680
GAACATAAAC CGGTTGGAAT TGGTCATTCT CACAGTAACT ATGAACTGTT TAAACCCGAA    1740
GAAGGAGTTG CTAAAAAAGA AGGGAATAAA GTTTATACTG GAGAAGAATT AACGAATGTT    1800
GTTAATTTGT TAAAAAATAG TACGTTTAAT AATCAAAACT TTACTCTAGC CAATGGTCAA    1860
AAACGCGTTT CTTTTAGTTT TCCGCCTGAA TTGGAGAAAA AATTAGGTAT CAATATGCTA    1920
GTAAAATTAA TAACACCAGA TGGAAAAGTA TTGGAGAAAG TATCTGGTAA AGTATTTGGA    1980
GAAGGAGTAG GGAATATTGC AAACTTTGAA TTAGATCAAC CTTATTTACC AGGACAAACA    2040
TTTAAGTATA CTATCGCTTC AAAAGATTAT CCAGAAGTAA GTTATGATGG TACATTTACA    2100
GTTCCAACCT CTTTAGCTTA CAAAATGGCC AGTCAAACGA TTTTCTATCC TTTCCATGCA    2160
GGGGATACTT ATTTAAGAGT GAACCCTCAA TTTGCAGTGC CTAAAGGAAC TGATGCTTTA    2220
GTCAGAGTGT TTGATGAATT TCATGGAAAT GCTTATTTAG AAAATAACTA TAAAGTTGGT    2280
GAAATCAAAT TACCGATTCC GAAATTAAAC CAAGGAACAA CCAGAACGGC CGGAAATAAA    2340
ATTCCTGTAA CCTTCATGGC AAATGCTTAT TTGGACAATC AATCGACTTA TATTGTGGAA    2400
GTACCTATCT TGGAAAAAGA AAATCAAACT GATAAACCAA GTATTCTACC ACAATTTAAA    2460
AGGAATAAAG CACAAGAAAA CTCAAAACTT GATGAAAAGG TAGAAGAACC AAAGACTAGT    2520
GAGAAGGTAG AAAAAGAAAA ACTTTCTGAA ACTGGGAATA GTACTAGTAA TTCAACGTTA    2580
GAAGAAGTTC CTACAGTGGA TCCTGTACAA GAAAAAGTAG CAAAATTTGC TGAAAGTTAT    2640
GGGATGAAGC TAGAAAATGT CTTGTTTAAT ATGGACGGAA CAATTGAATT ATATTTACCA    2700
TCAGGAGAAG TCATTAAAAA GAATATGGCA GATTTTACAG GAGAAGCACC TCAAGGAAAT    2760
GGTGAAAATA AACCATCTGA AAATGGAAAA GTATCTACTG GAACAGTTGA GAACCAACCA    2820
ACAGAAAATA AACCAGCAGA TTCTTTACCA GAGGCACCAA ACGAAAAACC TGTAAAACCA    2880
GAAAACTCAA CGGATAATGG AATGTTGAAT CCAGAAGGGA ATGTGGGGAG TGACCCTATG    2940
TTAGATCCAG CATTAGAGGA AGCTCCAGCA GTAGATCCTG TACAAGAAAA ATTAGAAAAA    3000
TTTACAGCTA GTTACGGATT AGGCTTAGAT AGTGTTATAT TCAATATGGA TGGAACGATT    3060
GAATTAAGAT TGCCAAGTGG AGAAGTGATA AAAAAGAATT TATCTGATTT CATAGCGTAA    3120
```

(SEQ ID NO: 1)

FIGURE 1

```
MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK      50
SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF     100
SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK     150
DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA     200
YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT     250
QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR     300
TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV     350
VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF     400
HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GTSHEKHEED GYGFDANRII     450
AEDESGFVMS HGDHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS     500
HEQDYPGNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH     550
GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV     600
VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV     650
LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT     700
VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN     750
AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE     800
VPILEKENQT DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE     850
TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP     900
SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP     950
EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK    1000
FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDFIA (SEQ ID NO: 2) 1039
```

FIGURE 2

```
ATGAAAATCA ATAAAAAATA TCTAGCTGGG TCAGTAGCTA CACTTGTTTT AAGTGTCTGT      60
GCTTATGAAC TAGGTTTGCA TCAAGCTCAA ACTGTAAAAG AAAATAATCG TGTTTCCTAT     120
ATAGATGGAA AACAAGCGAC GCAAAAAACG GAGAATTTGA CTCCTGATGA GGTTAGCAAG     180
CGTGAAGGAA TCAACGCCGA ACAAATCGTC ATCAAGATTA CGGATCAAGG TTATGTGACC     240
TCTCATGGAG ACCATTATCA TTACTATAAT GGCAAGGTCC CTTATGATGC CATCATCAGT     300
GAAGAGCTCC TCATGAAAGA TCCGAATTAT CAGTTGAAGG ATTCAGACAT TGTCAATGAA     360
ATCAAGGGTG GTTATGTCAT TAAGGTAAAC GGTAAATACT ATGTTTACCT TAAGGATGCA     420
GCTCATGCGG ATAATGTCCG TACAAAAGAA GAAATCAATC GGCAAAAACA AGAACATAGT     480
CAGCATCGTG AAGGAGGGAC TTCAGCAAAC GATGGTGCGG TAGCCTTTGC ACGTTCACAG     540
GGACGCTACA CCACAGATGA TGGTTATATC TTCAATGCAT CTGATATCAT CGAAGATACG     600
GGCGATGCCT ATATCGTTCC TCATGGAGAT CATTACCATT ACATTCCTAA GAATGAGTTA     660
TCAGCTAGCG AGTTGGCTGC TGCAGAAGCC TTCCTATCTG GTCGGGAAAA TCTGTCAAAT     720
TTAAGAACCT ATCGCCGACA AAATAGCGAT AACACTCCAA GAACAAACTG GGTACCTTCT     780
GTAAGCAATC CAGGAACTAC AAATACTAAC ACAAGCAACA ACAGCAACAC TAACAGTCAA     840
GCAAGTCAAA GTAATGACAT TGATAGTCTC TTGAAACAGC TCTACAAACT GCCTTTGAGT     900
CAACGCCATG TAGAATCTGA TGGCCTTATT TTCGACCCAG CGCAAATCAC AAGTCGAACC     960
GCCAGAGGTG TAGCTGTCCC TCATGGTAAC CATTACCACT TTATCCCTTA TGAACAAATG    1020
TCTGAATTGG AAAAACGAAT TGCTCGTATT ATTCCCCTTC GTTATCGTTC AAACCATTGG    1080
GTACCAGATT CAAGACCAGA AGAACCAAGT CCACAACCGA CTCCAGAACC TAGTCCAAGT    1140
CCGCAACCTG CACCAAATCC TCAACCAGCT CCAAGCAATC CAATTGATGA GAAATTGGTC    1200
AAAGAAGCTG TTCGAAAAGT AGGCGATGGT TATGTCTTTG AGGAGAATGG AGTTTCTCGT    1260
TATATCCCAG CCAAGAATCT TTCAGCAGAA ACAGCAGCAG GCATTGATAG CAAACTGGCC    1320
AAGCAGGAAA GTTTATCTCA TAAGCTAGGA GCTAAGAAAA CTGACCTCCC ATCTAGTGAT    1380
CGAGAATTTT ACAATAAGGC TTATGACTTA CTAGCAAGAA TTCACCAAGA TTTACTTGAT    1440
AATAAAGGTC GACAAGTTGA TTTTGAGGCT TTGGATAACC TGTTGGAACG ACTCAAGGAT    1500
GTCTCAAGTG ATAAAGTCAA GTTAGTGGAT GATATTCTTG CCTTCTTAGC TCCGATTCGT    1560
CATCCAGAAC GTTTAGGAAA ACCAAATGCG CAAATTACCT ACACTGATGA TGAGATTCAA    1620
GTAGCCAAGT TGGCAGGCAA GTACACAACA GAAGACGGTT ATATCTTTGA TCCTCGTGAT    1680
ATAACCAGTG ATGAGGGGGA TGCCTATGTA ACTCCACATA TGACCCATAG CCACTGGATT    1740
AAAAAAGATA GTTTGTCTGA AGCTGAGAGA GCGGCAGCCC AGGCTTATGC TAAAGAGAAA    1800
GGTTTGACCC CTCCTTCGAC AGACCATCAG GATTCAGGAA ATACTGAGGC AAAAGGAGCA    1860
GAAGCTATCT ACAACCGCGT GAAAGCAGCT AAGAAGGTGC CACTTGATCG TATGCCTTAC    1920
AATCTTCAAT ATACTGTAGA AGTCAAAAAC GGTAGTTTAA TCATACCTCA TTATGACCAT    1980
TACCATAACA TCAAATTTGA GTGGTTTGAC GAAGGCCTTT ATGAGGCACC TAAGGGGTAT    2040
ACTCTTGAGG ATCTTTTGGC GACTGTCAAG TACTATGTCG AACATCCAAA CGAACGTCCG    2100
CATTCAGATA ATGGTTTTGG TAACGCTAGC GACCATGTTC AAAGAAACAA AAATGGTCAA    2160
GCTGATACCA ATCAAACGGA AAAACCAAGC GAGGAGAAAC CTCAGACAGA AAAACCTGAG    2220
GAAGAAACCC CTCGAGAAGA GAAACCACAA AGCGAGAAAC CAGAGTCTCC AAAACCAACA    2280
GAGGAACCAG AAGAAGAATC ACCAGAGGAA TCAGAAGAAC CTCAGGTCGA GACTGAAAAG    2340
GTTGAAGAAA AACTGAGAGA GGCTGAAGAT TTACTTGGAA AAATCCAGGA TCCAATTATC    2400
AAGTCCAATG CCAAAGAGAC TCTCACAGGA TTAAAAAATA ATTTACTATT TGGCACCCAG    2460
GACAACAATA CTATTATGGC AGAAGCTGAA AAACTATTGG CTTTATTAAA GGAGAGTAAG    2520
TAA    (SEQ ID NO: 3)                                                2523
```

FIGURE 3

```
MKINKKYLAG SVATLVLSVC AYELGLHQAQ TVKENNRVSY IDGKQATQKT       50
ENLTPDEVSK REGINAEQIV IKITDQGYVT SHGDHYHYYN GKVPYDAIIS      100
EELLMKDPNY QLKDSDIVNE IKGGYVIKVN GKYYVYLKDA AHADNVRTKE      150
EINRQKQEHS QHREGGTSAN DGAVAFARSQ GRYTTDDGYI FNASDIIEDT      200
GDAYIVPHGD HYHYIPKNEL SASELAAAEA FLSGRENLSN LRTYRRQNSD      250
NTPRTNWVPS VSNPGTTNTN TSNNSNTNSQ ASQSNDIDSL LKQLYKLPLS      300
QRHVESDGLI FDPAQITSRT ARGVAVPHGN HYHFIPYEQM SELEKRIARI      350
IPLRYRSNHW VPDSRPEEPS PQPTPEPSPS PQPAPNPQPA PSNPIDEKLV      400
KEAVRKVGDG YVFEENGVSR YIPAKNLSAE TAAGIDSKLA KQESLSHKLG      450
AKKTDLPSSD REFYNKAYDL LARIHQDLLD NKGRQVDFEA LDNLLERLKD      500
VSSDKVKLVD DILAFLAPIR HPERLGKPNA QITYTDDEIQ VAKLAGKYTT      550
EDGYIFDPRD ITSDEGDAYV TPHMTHSHWI KKDSLSEAER AAAQAYAKEK      600
GLTPPSTDHQ DSGNTEAKGA EAIYNRVKAA KKVPLDRMPY NLQYTVEVKN      650
GSLIIPHYDH YHNIKFEWFD EGLYEAPKGY TLEDLLATVK YYVEHPNERP      700
HSDNGFGNAS DHVQRNKNGQ ADTNQTEKPS EEKPQTEKPE EETPREEKPQ      750
SEKPESPKPT EEPEEESPEE SEEPQVETEK VEEKLREAED LLGKIQDPII      800
KSNAKETLTG LKNNLLFGTQ DNNTIMAEAE KLLALLKESK (SEQ ID NO: 4)  840
```

FIGURE 4

```
ATGGAGAATA TAGACATGTT TAAATCAAAT CATGAGCGAA GAATGCGTTA TTCCATTCGT      60
AAATTTAGTG TAGGAGTAGC TAGCGTAGCT GTTGCCAGTC TTTTTATGGG AAGTGTTGTA     120
CATGCGACAG AGAAAGAGGG AAGTACCCAA GCAGCCACTT CTTTTAATAG GGGAAATGGA     180
AGTCAGGCAG AACAACGTGG AGAACTCGAT TTAGAACGAG ATAAGGCAAT GAAAGCGGTC     240
AGTGAATATG TAGGAAAAAT GGTGAGAGAT GCCTATGTAA AATCAGATAG AAAACGACAT     300
AAAAATACTG TAGCTCTAGT TAACCAGTTG GGAAACATTA AGAACAGGTA TTTGAATGAA     360
ATAGTTCATT CAACCTCAAA AAGCCAACTA CAGGAACTGA TGATGAAGAG TCAATCAGAA     420
GTAGATGAAG CTGTGTCTAA ATTTGAAAAG GACTCATTTT CTTCGTCAAG TTCAGGATCC     480
TCCACTAAAC CAGAAACTCC GCAGCCGGAA AATCCAGAGC ATCAAAAACC AACAACTCCA     540
TCTCCGGATA CCAAACCAAG CCCTCAACCA GAAGGCAAGA AACCAAGCGT ACCAGACTCA     600
AATCAGGAAA AAGAAAAAGC TAAGCTTGCT GTAGTAACCT ACATGAGCAA GATTTTAGAT     660
GATATACAAA AACATCATCT GCAGAAAGAA AAACATCGTC AGATTGTTGC TCTTATTAAG     720
GAGCTTGATG AGCTTAAAAA GCAAGCTCTT TCTGAAATTG ATAATGTAAA TACCAAAGTA     780
GAAATTGAAA ATACAGTCCA CAAGATATTT GCAGACATGG ATGCAGTTGT GACTAAATTC     840
AAAAAAGGCT TAACTCAGGA CACACCAAAA GAACCAGGTA ACAAAAAACC ATCTGCTCCA     900
AAACCAGGTA TGCAACCAAG TCCTCAACCA GAGGTTAAAC CGCAGCTGGA AAAACCAAAA     960
CCAGAGGTTA AACCGCAACC AGAAAAACCA AAACCAGAGG TTAAACCGCA GCCGGAAAAA    1020
CCAAAACCAG AGGTTAAACC GCAGCCGGAA AAACCAAAAC CAGAGGTTAA ACCGCAGCCG    1080
GAAAAACCAA AACCAGAGGT TAAACCGCAG CCGGAAAAAC CAAAACCAGA GGTTAAACCG    1140
CAGCCGGAAA AACCAAAACC AGAGGTTAAA CCGCAGCCGG AAAAACCAAA ACCAGAGGTT    1200
AAACCGCAGC CGGAAAAACC AAAACCAGAG GTTAAACCGC AGCCGGAAAA ACCAAAACCA    1260
GAGGTTAAAC CGCAGCCGGA AAAACCAAAA CCAGAGGTTA AACCGCAACC AGAAAAACCA    1320
AAACCAGAGG TTAAACCGCA ACCAGAAAAA CCAAAACCAG ATAATAGCAA GCCACAAGCA    1380
GATGATAAGA AGCCATCAAC TACAAATAAT TTAAGCAAGG ACAAGCAACC TTCTAACCAA    1440
GCTTCAACAA ACGAAAAAGC AACAAATAAA CCGAAGAAGT CATTGCCATC AACTGGATCT    1500
ATTTCAAATC TAGCACTTGA AATTGCAGGT CTTCTTACCT TGGCGGGGGC AACCATTCTT    1560
GCTAAGAAAA GAATGAAATA G  (SEQ ID NO: 5)                              1581
```

FIGURE 5

```
MENIDMFKSN HERRMRYSIR KFSVGVASVA VASLFMGSVV HATEKEGSTQ      50
AATSFNRGNG SQAEQRGELD LERDKAMKAV SEYVGKMVRD AYVKSDRKRH     100
KNTVALVNQL GNIKNRYLNE IVHSTSKSQL QELMMKSQSE VDEAVSKFEK     150
DSFSSSSSGS STKPETPQPE NPEHQKPTTP SPDTKPSPQP EGKKPSVPDI     200
NQEKEKAKLA VVTYMSKILD DIQKHHLQKE KHRQIVALIK ELDELKKQAL     250
SEIDNVNTKV EIENTVHKIF ADMDAVVTKP KKGLTQDTPK EPGNKKPSAP     300
KPGMQPSPQP EVKPQLEKPK PEVKPQPEKP KPEVKPQPEK PKPEVKPQPE     350
KPKPEVKPQP EKPKPEVKPQ PEKPKPEVKP QPEKPKPEVK PQPEKPKPEV     400
KPQPEKPKPE VKPQPEKPKP EVKPQPEKPK PEVKPQPEKP KPEVKPQPEK     450
PKPDNSKPQA DDKKPSTTNN LSKDKQPSNQ ASTNEKATNK PKKSLPSTGS     500
ISNLALEIAG LLTLAGATIL AKKRMK      (SEQ ID NO: 6)           526
```

FIGURE 6

```
ATGAAATTTA GTAAAAAATA TATAGCAGCT GGATCAGCTG TTATCGTATC CTTGAGTCTA      60
TGTGCCTATG CACTAAACCA GCATCGTTCG CAGGAAAATA AGGACAATAA TCGTGTCTCT     120
TATGTGGATG GCAGCCAGTC AAGTCAGAAA AGTGAAAACT TGACACCAGA CCAGGTTAGC     180
CAGAAAGAAG GAATTCAGGC TGAGCAAATT GTAATCAAAA TTACAGATCA GGGCTATGTA     240
ACGTCACACG GTGACCACTA TCATTACTAT AATGGGAAAG TTCCTTATGA TGCCCTCTTT     300
AGTGAAGAAC TCTTGATGAA GGATCCAAAC TATCAACTTA AAGACGCTGA TATTGTCAAT     360
GAAGTCAAGG GTGGTTATAT CATCAAGGTC GATGGAAAAT ATTATGTCTA CCTGAAAGAT     420
GCAGCTCATG CTGATAATGT TCGAACTAAA GATGAAATCA ATCGTCAAAA ACAAGAACAT     480
GTCAAAGATA ATGAGAAGGT TAACTCTAAT GTTGCTGTAG CAAGGTCTCA GGGACGATAT     540
ACGACAAATG ATGGTTATGT CTTTAATCCA GCTGATATTA TCGAAGATAC GGGTAATGCT     600
TATATCGTTC CTCATGGAGG TCACTATCAC TACATTCCCA AAAGCGATTT ATCTGCTAGT     660
GAATTAGCAG CAGCTAAAGC ACATCTGGCT GGAAAAAATA TGCAACCGAG TCAGTTAAGC     720
TATTCTTCAA CAGCTAGTGA CAATAACACG CAATCTGTAG CAAAAGGATC AACTAGCAAG     780
CCAGCAAATA AATCTGAAAA TCTCCAGAGT CTTTTGAAGG AACTCTATGA TTCACCTAGC     840
GCCCAACGTT ACAGTGAATC AGATGGCCTG GTCTTTGACC CTGCTAAGAT TATCAGTCGT     900
ACACCAAATG GAGTTGCGAT TCCGCATGGC GACCATTACC ACTTTATTCC TTACAGCAAG     960
CTTTCTGCTT TAGAAGAAAA GATTGCCAGA ATGGTGCCTA TCAGTGGAAC TGGTTCTACA    1020
GTTTCTACAA ATGCAAAACC TAATGAAGTA GTGTCTAGTC TAGGCAGTCT TTCAAGCAAT    1080
CCTTCTTCTT TAACGACAAG TAAGGAGCTC TCTTCAGCAT CTGATGGTTA TATTTTTAAT    1140
CCAAAAGATA TCGTTGAAGA AACGGCTACA GCTTATATTG TAAGACATGG TGATCATTTC    1200
CATTACATTC CAAAATCAAA TCAAATTGGG CAACCGACTC TTCCAAACAA TAGTCTAGCA    1260
ACACCTTCTC CATCTCTTCC AATCAATCCA GGAACTTCAC ATGAGAAACA TGAAGAAGAT    1320
GGATACGGAT TTGATGCTAA TCGTATTATC GCTGAAGATG AATCAGGTTT TGTCATGAGT    1380
CACGGAGACC ACAATCATTA TTTCTTCAAG AAGGACTTGA CAGAAGAGCA AATTAAGGTG    1440
CGCAAAAACA TTTAG    (SEQ ID NO: 7)                                    1455
```

FIGURE 7

```
MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK      50
SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF     100
SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK     150
DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA     200
YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT     250
QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR     300
TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV     350
VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF     400
HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GTSHEKHEED GYGFDANRII     450
AEDESGFVMS HGDHNHYFFK KDLTEEQIKV RKNI    (SEQ ID NO: 8)     484
```

FIGURE 8

```
ATGAAAGATT TAGATAAAAA AATCGAAGAA AAAATTGCTG GCATTATGAA ACAATATGGT      60
GTCAAACGTG AAAGTATTGT CGTGAATAAA GAAAAAAATG CGATTATTTA TCCGCATGGA     120
GATCACCATC ATGCAGATCC GATTGATGAA CATAAACCGG TTGGAATTGG TCATTCTCAC     180
AGTAACTATG AACTGTTTAA ACCCGAAGAA GGAGTTGCTA AAAAAGAAGG GAATAAAGTT     240
TATACTGGAG AAGAATTAAC GAATGTTGTT AATTTGTTAA AAAATAGTAC GTTTAATAAT     300
CAAAACTTTA CTCTAGCCAA TGGTCAAAAA CGCGTTTCTT TTAGTTTTCC GCCTGAATTG     360
GAGAAAAAAT TAGGTATCAA TATGCTAGTA AAATTAATAA CACCAGATGG AAAAGTATTG     420
GAGAAAGTAT CTGGTAAAGT ATTTGGAGAA GGAGTAGGGA ATATTGCAAA CTTTGAATTA     480
GATCAACCTT ATTTACCAGG ACAAACATTT AAGTATACTA TCGCTTCAAA AGATTATCCA     540
GAAGTAAGTT ATGATGGTAC ATTTACAGTT CCAACCTCTT TAGCTTACAA AATGGCCAGT     600
CAAACGATTT TCTATCCTTT CCATGCAGGG GATACTTATT TAAGAGTGAA CCCTCAATTT     660
GCAGTGCCTA AAGGAACTGA TGCTTTAGTC AGAGTGTTTG ATGAATTTCA TGGAAATGCT     720
TATTTAGAAA ATAACTATAA AGTTGGTGAA ATCAAATTAC CGATTCCGAA ATTAAACCAA     780
GGAACAACCA GAACGGCCGG AAATAAAATT CCTGTAACCT TCATGGCAAA TGCTTATTTG     840
GACAATCAAT CGACTTATAT TGTGGAAGTA CCTATCTTGG AAAAAGAAAA TCAAACTGAT     900
AAACCAAGTA TTCTACCACA ATTTAAAAGG AATAAAGCAC AAGAAAACTC AAAACTTGAT     960
GAAAAGGTAG AAGAACCAAA GACTAGTGAG AAGGTAGAAA AAGAAAAACT TTCTGAAACT    1020
GGGAATAGTA CTAGTAATTC AACGTTAGAA GAAGTTCCTA CAGTGGATCC TGTACAAGAA    1080
AAAGTAGCAA AATTTGCTGA AAGTTATGGG ATGAAGCTAG AAAATGTCTT GTTTAATATG    1140
GACGGAACAA TTGAATTATA TTTACCATCA GGAGAAGTCA TTAAAAAGAA TATGGCAGAT    1200
TTTACAGGAG AAGCACCTCA AGGAAATGGT GAAAATAAAC CATCTGAAAA TGGAAAAGTA    1260
TCTACTGGAA CAGTTGAGAA CCAACCAACA GAAAATAAAC CAGCAGATTC TTTACCAGAG    1320
GCACCAAACG AAAAACCTGT AAAACCAGAA AACTCAACGG ATAATGGAAT GTTGAATCCA    1380
GAAGGGAATG TGGGGAGTGA CCCTATGTTA GATCCAGCAT TAGAGGAAGC TCCAGCAGTA    1440
GATCCTGTAC AAGAAAAATT AGAAAAATTT ACAGCTAGTT ACGGATTAGG CTTAGATAGT    1500
GTTATATTCA ATATGGATGG AACGATTGAA TTAAGATTGC CAAGTGGAGA AGTGATAAAA    1560
AAGAATTTAT CTGATTTCAT AGCGTAA     (SEQ ID NO: 9)                     1587
```

FIGURE 9

```
MKDLDKKIEE KIAGIMKQYG VKRESIVVNK EKNAIIYPHG DHHHADPIDE      50
HKPVGIGHSH SNYELFKPEE GVAKKEGNKV YTGEELTNVV NLLKNSTFNN     100
QNFTLANGQK RVSFSFPPEL EKKLGINMLV KLITPDGKVL EKVSGKVFGE     150
GVGNIANFEL DQPYLPGQTF KYTIASKDYP EVSYDGTFTV PTSLAYKMAS     200
QTIFYPFHAG DTYLRVNPQF AVPKGTDALV RVFDEFHGNA YLENNYKVGE     250
IKLPIPKLNQ GTTRTAGNKI PVTFMANAYL DNQSTYIVEV PILEKENQTD     300
KPSILPQFKR NKAQENSKLD EKVEEPKTSE KVEKEKLSET GNSTSNSTLE     350
EVPTVDPVQE KVAKFAESYG MKLENVLFNM DGTIELYLPS GEVIKKNMAD     400
FTGEAPQGNG ENKPSENGKV STGTVENQPT ENKPADSLPE APNEKPVKPE     450
NSTDNGMLNP EGNVGSDPML DPALEEAPAV DPVQEKLEKF TASYGLGLDS     500
VIFNMDGTIE LRLPSGEVIK KNLSDFIA  (SEQ ID NO: 10)            528
```

FIGURE 10

```
BVH3 WU2          1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
BVH3 RX1          1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
BVH3 JNR7/87      1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
BVH3 SP64         1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
BVH3 P4241        1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
BVH3 A66          1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV   60
                    ************************************************************

BVH3 WU2         61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
BVH3 RX1         61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
BVH3 JNR7/87     61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
BVH3 SP64        61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
BVH3 P4241       61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
BVH3 A66         61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD  120
                    ************************************************************

BVH3 WU2        121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
BVH3 RX1        121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
BVH3 JNR7/87    121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
BVH3 SP64       121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
BVH3 P4241      121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
BVH3 A66        121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA  180
                    ************************************************************

BVH3 WU2        181 YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
BVH3 RX1        181 YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
BVH3 JNR7/87    181 YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
BVH3 SP64       181 YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
BVH3 P4241      181 YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
BVH3 A66        181 YIVPHRGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK  240
                    ***** ********************************************************

BVH3 WU2        241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
BVH3 RX1        241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
BVH3 JNR7/87    241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
BVH3 SP64       241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
BVH3 P4241      241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
BVH3 A66        241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK  300
                    ************************************************************

BVH3 WU2        301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
BVH3 RX1        301 LSALEEKIARRVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
BVH3 JNR7/87    301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
BVH3 SP64       301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
BVH3 P4241      301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
BVH3 A66        301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN  360
                    ********** ************************************************

BVH3 WU2        361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED  420
BVH3 RX1        361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGISHEKHEED  420
BVH3 JNR7/87    361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED  420
BVH3 SP64       361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED  420
BVH3 P4241      361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED  420
BVH3 A66        361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED  420
                    ************************************************************ ********

BVH3 WU2        421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
BVH3 RX1        421 GYGFDANRIIAEDESGFIMSHGNHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
BVH3 JNR7/87    421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
BVH3 SP64       421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
BVH3 P4241      421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
BVH3 A66        421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS  480
                    ****************.**** ****************************************
```

```
BVH3 WU2        481 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
BVH3 RX1        481 HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
BVH3 JNR7/87    481 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
BVH3 SP64       481 HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
BVH3 P4241      481 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
BVH3 A66        481 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID  540
                    ****** **************************************************

BVH3 WU2        541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
BVH3 RX1        541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
BVH3 JNR7/87    541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
BVH3 SP64       541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
BVH3 P4241      541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
BVH3 A66        541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ  600
                    ************************************************************

BVH3 WU2        601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
BVH3 RX1        601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
BVH3 JNR7/87    601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
BVH3 SP64       601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
BVH3 P4241      601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
BVH3 A66        601 KRVSFSFPPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT  660
                    ************************************************************

BVH3 WU2        661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3 RX1        661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3 JNR7/87    661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3 SP64       661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3 P4241      661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3 A66        661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
                    ************************************************************

BVH3 WU2        721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3 RX1        721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3 JNR7/87    721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3 SP64       721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3 P4241      721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3 A66        721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
                    ************************************************************

BVH3 WU2        781 VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3 RX1        781 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3 JNR7/87    781 VPILEKENQTDKPSILPQFKRNKAQENLKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3 SP64       781 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3 P4241      781 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3 A66        781 VPILEKENQTDKPSILPQFKRNKAQENSKFDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
                    ***************************** * ******************************

BVH3 WU2        841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3 RX1        841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3 JNR7/87    841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3 SP64       841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3 P4241      841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3 A66        841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
                    ************************************************************

BVH3 WU2        901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3 RX1        901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3 JNR7/87    901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3 SP64       901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3 P4241      901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3 A66        901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
                    ************************************************************

BVH3 WU2        961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
BVH3 RX1        961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
BVH3 JNR7/87    961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
BVH3 SP64       961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
BVH3 P4241      961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDFIA 1019
BVH3 A66        961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDLIA 1019
                    ************************************************************ **
```

FIGURE 11

```
BVH11-2 SP64      1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11-2 JNR7/87   1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11-2 P4241     1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11-2 A66       1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11-2 WU2       1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11-2 Rx1       1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 P4241       1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 WU2         1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 A66         1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 Rx1         1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 JNR7/87     1 CSYELGRHQAGQDKKESNRVAYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11 SP63        1 CSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEVSKREGINAEQIVIKITDQGY  60
BVH11  SP64       1 CAYELGLHQA-QTVKENNRVSYIDGKQATQKTENLTPDEVSKREGINAEQIVIKITDQGY  59
                    *.**** *** *  ** ***.**** ** **.*****************************

BVH11-2 SP64     61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120
BVH11-2 JNR7/87  61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120
BVH11-2 P4241    61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11-2 A66      61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11-2 WU2      61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11-2 Rx1      61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11 P4241      61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120
BVH11 WU2        61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYGYLK 120
BVH11 A66        61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11 Rx1        61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120
BVH11 JNR7/87    61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 120
BVH11 SP63       61 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLK 120
BVH11  SP64      60 VTSHGDHYHYYYNGKVPYDAIISEELLMKDPNYQLKDSDIVNEIKGGYVIKVNGKYYVYLK 119
                   ********************************************************* **** ***

BVH11-2 SP64    121 DAAHADNIRTKEEIKRQKQEHSHNHNSRA---DNAVAAARAQGRYTTDDGYIFNASDIIE 177
BVH11-2 JNR7/87 121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11-2 P4241   121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11-2 A66     121 DAAHADNIRTKEEIKRQRQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11-2 WU2     121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11-2 Rx1     121 DAAHADNIRTKEEIKRQKQERSHNHNSRA---DNAVAAARAQGRYTTDDGYIFNASDIIE 177
BVH11 P4241     121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11 WU2       121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11 A66       121 DAAHADNIRTKEEIKRQKQEHSHNHGGGSN--DQAVVAARAQGRYTTDDGYIFNASDIIE 178
BVH11 Rx1       121 DAAHADNIRTKEEIKRQKQERSHNHNSRA---DNAVAAARAQGRYTTDDGYIFNASDIIE 177
BVH11 JNR7/87   121 DAAHADNIRTKEEIKRQKQERSHNHNSRA---DNAVAAARAQGRYTTDDGYIFNASDIIE 177
BVH11 SP63      121 DAAHADNIRTKEEIKRQKQERSHNHNSRA---DNAVAAARAQGRYTTDDGYIFNASDIIE 177
BVH11  SP64     120 DAAHADNVRTKEEINRQKQEHSQHREGGTSANDGAVAFARSQGRYTTDDGYIFNASDIIE 179
                   *******.******* **.**.*...   .   * **  **.********************

BVH11-2 SP64    178 DTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPVQPRLSEN 237
BVH11-2 JNR7/87 179 DTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPAQPRLSEN 238
BVH11-2 P4241   179 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11-2 A66     179 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11-2 WU2     179 DTGDAYIVPRGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11-2 Rx1     178 DTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 237
BVH11 P4241     179 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11 WU2       179 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11 A66       179 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 238
BVH11 Rx1       178 DTGDAYIVPHGDHYHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 237
BVH11 JNR7/87   178 DTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSSSSYNANPAQPRLSEN 237
BVH11 SP63      178 DTGDAYIVPHGNHFHYIPKSDLSASELAAAQAYWNGKQGSRPSSSSSHNANPAQPRLSEN 237
BVH11  SP64     180 DTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRENLSNLRTYRRQNSDNTPRTNWV 239
                   *********.* *.***** .*********.*.  *...    .   .      **
```

```
BVH11-2 SP64     238 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 285
BVH11-2 JNR7/87  239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11-2 P4241    239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11-2 A66      239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11-2 WU2      239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11-2 Rx1      238 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERRVESDGLIFDPAQITS 286
BVH11 P4241      239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 285
BVH11 WU2        239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11 A66        239 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11 Rx1        238 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 286
BVH11 JNR7/87    238 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 285
BVH11 SP63       238 HNLTVTPTYHQN------------QGENISSLLRELYAKPLSERHVESDGLIFDPAQITS 285
BVH11  SP64      240 PSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLPLSQRHVESDGLIFDPAQITS 299
                     ..   * . *             *   * *** . ** *** . * . ***************

BVH11-2 SP64     286 RTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 345
BVH11-2 JNR7/87  287 RTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 346
BVH11-2 P4241    287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQ----PS 342
BVH11-2 A66      287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQ----PS 342
BVH11-2 WU2      287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQ----PS 342
BVH11-2 Rx1      286 RTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 345
BVH11 P4241      287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 345
BVH11 WU2        287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQ----PS 342
BVH11 A66        287 RTARGVAVPHGNHYHFIPYEQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQ----PS 342
BVH11 Rx1        286 RTANGVAVPHGDHYHFIPYSQLSPLEEKLARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 345
BVH11 JNR7/87    286 RTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEEPSPQPTPEPS 345
BVH11 SP63       286 RTARGVAVPHGNHYHFIPYSQMSELEERIARIIPLRYRSNHWVPDSRPEQPSPQSTPEPS 345
BVH11  SP64      300 RTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSNHWVPDSRPEEPSPQPTPEPS 359
                     *** ******* ******* *.* ** ..********************.****    **

BVH11-2 SP64     346 PSLQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 405
BVH11-2 JNR7/87  347 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 406
BVH11-2 P4241    343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11-2 A66      343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11-2 WU2      343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11-2 Rx1      346 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11 P4241      343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 405
BVH11 WU2        343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11 A66        343 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSK 402
BVH11 Rx1        346 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVPRYIPAKDLSAETAAGIDSK 405
BVH11 JNR7/87    346 PSP------QPAPSNPIDEKLVKEAVRKVGDGYVFEENGVPRYIPAKDLSAETAAGIDSK 405
BVH11 SP63       346 PSPQSAPNPQPAPSNPIDEKLVKEVVRKVGDGYVFEKNGVSRYIPAKNLSAETAAGIDSK 399
BVH11  SP64      360 PSPQPAPNPQPAPSNPIDEKLVKEAVRKVGDGYVFEENGVSRYIPAKNLSAETAAGIDSK 419
                     **       *************** *********** *** ****** ************

BVH11-2 SP64     406 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEVLDNLLERL 465
BVH11-2 JNR7/87  407 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 466
BVH11-2 P4241    403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11-2 A66      403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11-2 WU2      403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11-2 Rx1      406 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 465
BVH11 P4241      403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11 WU2        403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11 A66        403 LAKQESLSHKLGTKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 462
BVH11 Rx1        406 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 465
BVH11 JNR7/87    400 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 459
BVH11 SP63       406 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 465
BVH11  SP64      420 LAKQESLSHKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERL 479
                     ************.****************************************** ********
```

```
BVH11-2 SP64     466 KDVSSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 525
BVH11-2 JNR7/87  467 KDVPSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 526
BVH11-2 P4241    463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11-2 A66      463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11-2 WU2      463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11-2 Rx1      466 KDVSSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 525
BVH11 P4241      463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11 WU2        463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11 A66        463 KDVSSDKVKLVEDILAFLAPIRHPERLGKPNSQITYTDDEIQVAKLAGKYTTEDGYIFDP 522
BVH11 Rx1        466 KDVSSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 525
BVH11 JNR7/87    460 KDVSSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 519
BVH11 SP63       466 EDVPSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 525
BVH11  SP64      480 KDVSSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP 539
                     ** ********.******************.*****************************

BVH11-2 SP64     526 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 585
BVH11-2 JNR7/87  527 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 586
BVH11-2 P4241    523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHRDSGNTEAK 582
BVH11-2 A66      523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 582
BVH11-2 WU2      523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 582
BVH11-2 Rx1      526 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 585
BVH11 P4241      523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 582
BVH11 WU2        523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 582
BVH11 A66        523 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 582
BVH11 Rx1        526 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 585
BVH11 JNR7/87    520 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 579
BVH11 SP63       526 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 585
BVH11  SP64      540 RDITSDEGDAYVTPHMTHSHWIKKDSLSEAERAAAQAYAKEKGLTPPSTDHQDSGNTEAK 599
                     ********************************************************.********

BVH11-2 SP64     586 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 645
BVH11-2 JNR7/87  587 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 646
BVH11-2 P4241    583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11-2 A66      583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11-2 WU2      583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11-2 Rx1      586 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 645
BVH11 P4241      583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11 WU2        583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11 A66        583 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 642
BVH11 Rx1        586 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 645
BVH11 JNR7/87    580 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 639
BVH11 SP63       586 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 645
BVH11  SP64      600 GAEAIYNRVKAAKKVPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK 659
                     ************************************************************

BVH11-2 SP64     646 GYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 690
BVH11-2 JNR7/87  647 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------VDQDSK 691
BVH11-2 P4241    643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11-2 A66      643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11-2 WU2      643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11-2 Rx1      646 GYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADTNQTEKPNEEKPQTEK 705
BVH11 P4241      643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11 WU2        643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11 A66        643 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVRKNK---------------ADQDSK 687
BVH11 Rx1        646 GYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNK-----------------NGQ 687
BVH11 JNR7/87    640 GYSLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNK-----------------NGQ 681
BVH11 SP63       646 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNK-----------------NGQ 687
BVH11  SP64      660 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNK-----------------NGQ 701
                     **.*********************************..**
```

```
BVH11-2 SP64     691 PDEDKEHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 750
BVH11-2 JNR7/87  692 PDEDKEHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 751
BVH11-2 P4241    688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11-2 A66      688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11-2 WU2      688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11-2 Rx1      706 PEEDKEHDEVSEPTHPESDEKENHVGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 765
BVH11 P4241      688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11 WU2        688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11 A66        688 PDEDKGHDEVSEPTHPESDEKENHAGLNPSADNLYKPSTDTEETEEEAEDTTDEAEIPQV 747
BVH11 Rx1        688 ADTNQTEKPNEEKPQTEKPEEETPREEKPQSEKPESPKPTEEPEEESPEESPEESEEPQV 747
BVH11 JNR7/87    682 ADTNQTEKPNEEKPQTEKPEEETPREEKPQSEKPESPKPTEEPEEESPEESPEESEEPQV 741
BVH11 SP63       688 ADTNQTEKPSEEKPQTEKPEEETPREEKPQSEKPESP----KPTEEPEEESPEESEEPQV 743
BVH11  SP64      702 ADTNQTEKPSEEKPQTEKPEEETPREEKPQSEKPESP----KPTEEPEEESPEESEEPQV 757
                     . .    * . * * *.    * ..   *       **  *.. .*.* ***

BVH11-2 SP64     751 ENSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 810
BVH11-2 JNR7/87  752 ENSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 811
BVH11-2 P4241    748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11-2 A66      748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11-2 WU2      748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11-2 Rx1      766 EYSVINAKIAEAEALLEKVTDSSIRQNAVETLTGLKSSLLLGTKDNNTISAEVDSLLALL 825
BVH11 P4241      748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11 WU2        748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11 A66        748 EHSVINAKIADAEALLEKVTDPSIRQNAMETLTGLKSSLLLGTKDNNTISAEVDSLLALL 807
BVH11 Rx1        748 ETEKVKEKLREAEDLLGKIQNPIIKSNAKETLTGLKNNLLFGTQDNNTIMAEAEKLLALL 807
BVH11 JNR7/87    742 ETEKVKEKLREAEDLLGKIQNPIIKSNAKETLTGLKNNLLFGTQDNNTIMAEAEKLLALL 801
BVH11 SP63       744 ETEKVEEKLREAEDLLGKIQDPIIKSNAKETLTGLKNNLLFGTQDNNTIMAEAEKLLALL 803
BVH11  SP64      758 ETEKVEEKLREAEDLLGKIQDPIIKSNAKETLTGLKNNLLFGTQDNNTIMAEAEKLLALL 817
                     *   . *. .** ** *.    *. ** *******  ** **.***** ** . *****

BVH11-2 SP64     811 KESQPAPIQ 819
BVH11-2 JNR7/87  812 KESQPAPIQ 820
BVH11-2 P4241    808 KKSQPAPIQ 816
BVH11-2 A66      808 KKSQPAPIQ 816
BVH11-2 WU2      808 KKSQPAPIQ 816
BVH11-2 Rx1      826 KESQPAPIQ 834
BVH11 P4241      808 KESK      811
BVH11 WU2        808 KESK      811
BVH11 A66        808 KESK      811
BVH11 Rx1        808 KESK      811
BVH11 JNR7/87    802 KESK      805
BVH11 SP63       804 KESK      807
BVH11  SP64      818 KESK      821
                     * *.
```

FIGURE 12

| BVH11-2 SP64 | BVH11 SP63 | BVH11 JNR.7/87 | BVH11-2 JNR.7/87 | BVH11 WU2 | BVH11-2 WU2 | BVH11 A66 | BVH11-2 A66 | BVH11 P4241 | BVH11-2 P4241 | BVH11 Rx-1 | BVH11-2 Rx-1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I 81% S 86% | I 88% S 90% | I 88% S 91% | I 82% S 87% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 80% S 85% | I 88% S 91% | I 81% S 85% | BVH11 SP64 |
| | I 87% S 90% | I 87% S 90% | I 98% S 98% | I 95% S 96% | I 96% S 97% | I 95% S 96% | I 96% S 97% | I 95% S 96% | I 96% S 97% | I 87% S 90% | I 94% S 95% | BVH11-2 SP64 |
| | | I 96% S 96% | I 88% S 91% | I 88% S 91% | I 87% S 90% | I 88% S 91% | I 87% S 90% | I 88% S 91% | I 87% S 90% | I 97% S 97% | I 89% S 91% | BVH11 SP63 |
| | | | I 87% S 90% | I 96% S 97% | I 86% S 90% | I 87% S 91% | I 86% S 90% | I 87% S 91% | I 86% S 90% | I 96% S 96% | I 88% S 90% | BVH11 JNR.7/87 |
| | | | | I 96% S 97% | I 97% S 98% | I 96% S 97% | I 97% S 98% | I 96% S 97% | I 97% S 98% | I 87% S 90% | I 94% S 95% | BVH11-2 JNR.7/87 |
| | | | | I 92% S 94% | I 98% S 98% | I 92% S 94% | I 98% S 98% | I 99% S 99% | I 98% S 98% | I 87% S 91% | I 92% S 94% | BVH11 WU2 |
| | | | | | I 98% S 98% | I 98% S 98% | I 99% S 99% | I 98% S 98% | I 99% S 99% | I 86% S 90% | I 93% S 95% | BVH11-2 WU2 |
| | | | | | | I 99% S 98% | I 99% S 99% | I 100% | I 99% S 99% | I 87% S 91% | I 92% S 94% | BVH11 A66 |
| | | | | | | | I 99% S 99% | I 99% S 99% | I 99% S 99% | I 86% S 90% | I 93% S 95% | BVH11-2 A66 |
| | | | | | | | | I 99% S 99% | I 99% S 99% | I 87% S 91% | I 92% S 94% | BVH11 P4241 |
| | | | | | | | | | | I 86% S 90% | I 93% S 95% | BVH11-2 P4241 |
| | | | | | | | | | | | I 91% S 92% | BVH11 Rx-1 |

FIGURE 13

```
AATTCCTTGT CGGGTAAGTT CCGACCCGCA CGAAAGGCGT AATGATTTGG GCACTGTCTC      60
AACGAGAGAC TCGGTGAAAT TTTAGTACCT GTGAAGATGC AGGTTACCCG CGACAGGACG     120
GAAAGACCCC ATGGAGCTTT ACTGCAGTTT GATATTGAGT GTCTGTACCA CATGTACAGG     180
ATAGGTAGGA GTCTAAGAGA TCGGGACGCC AGTTTCGAAG GAGACGCTGT TGGGATACTA     240
CCCTTGTGTT ATGGCCACTC TAACCCAGAT AGGTGATCCC TATCGGAGAC AGTGTCTGAC     300
GGGCAGTTTG ACTGGGGCGG TCGCCTCCTA AAAGGTAACG GAGGCGCCCA AAGGTTCCCT     360
CAGAATGGTT GGAAATCATT CGCAGAGTGT AAAGGTATAA GGGAGCTTGA CTGCGAGAGC     420
TACAACTCGA GCAGGGACGA AAGTCGGGCT TAGTGATCCG GTGGTTCCGT ATGGAAGGGC     480
CATCGCTCAA CGGATAAAAG CTACCCTGGG GATAACAGGC TTATCTCCCC CAAGAGTTCA     540
CATCGACGGG GAGGTTTGGC ACCTCGATGT CGGCTCGTCG CATCCTGGGG CTGTAGTCGG     600
TCCCAAGGGT TGGGCTGTTC GCCCATTAAA GCGGCACGCG AGCTGGGTTC AGAACGTCGT     660
GAGACAGTTC GGTCCCTATC CGTCGCGGGC GTAGGAAATT TGAGAGGATC TGCTCCTAGT     720
ACGAGAGGAC CAGAGTGGAC TTACCGCTGG TGTACCAGTT GTCTTGCCAA AGGCATCGCT     780
GGGTAGCTAT GTAGGGAAGG GATAAACGCT GAAAGCATCT AAGTGTGAAA CCCACCTCAA     840
GATGAGATTT CCCATGATTA TATATCAGTA AGAGCCCTGA GAGATGATCA GGTAGATAGG     900
TTAGAAGTGG AAGTGTGGCG ACACATGTAG CGGACTAATA CTAATAGCTC GAGGACTTAT     960
CCAAAGTAAC TGAGAATATG AAAGCGAACG GTTTTCTTAA ATTGAATAGA TATTCAATTT    1020
TGAGTAGGTA TTACTCAGAG TTAAGTGACG ATAGCCTAGG AGATACACCT GTACCCATGC    1080
CGAACACAGA AGTTAAGCCC TAGAACGCCG GAAGTAGTTG GGGGTTGCCC CCTGTGAGAT    1140
AGGGAAGTCG CTTAGCTCTA GGGAGTTTAG CTCAGCTGGG AGAGCATCTG CCTTACAAGC    1200
AGAGGGTCAG CGGTTCGATC CCGTTAACTC CCAAAGGTCC CGTAGTGTAG CGGTTATCAC    1260
GTCGCCCTGT CACGGCGAAG ATCGCGGGTT CGATTCCCGT CGGGACCGTT TAAGGTAACG    1320
CAAGTTATTT TAGACTCGTT AGCTCAGTTG GTAGAGCAAT TGACTTTTAA TCAATGGGTC    1380
ACTGGTTCGA GCCCAGTACG GGTCATATAT GCGGGTTTGG CGGAATTCTA ATCTCTTTGA    1440
AATCATCTTC TCTCACTTTC CAAAACTCTA TTACCTCTTA TTATACCACA TTTCAATCTT    1500
CAACTTCCCA GTAATATAAG CACCTCTGGC GAAAGAAGTT TCAATGTCCT AAAGTAATAA    1560
GTGAATCCAA TTCAGGAACT CCAAGAACAA AAGAAACATC TGGTGTCACA AGTATTGGAT    1620
GGCACAGAGT CACGTGGTAG TCTGACCCTA GCAGAAATTT TAAATAGTAA ACTATTTACT    1680
GGTTAATTAA ATGGTTAAAT AACCGGTTTA GAAAACTATT TAATAAAGTA AAAGAAGTTG    1740
AGAAAAAACT TCATCATTTA TTGAAATGAG GGATTTATGA AATTTAGTAA AAAATATATA    1800
GCAGCTGGAT CAGCTGTTAT CGTATCCTTG AGTCTATGTG CCTATGCACT AAACCAGCAT    1860
CGTTCGCAGG AAAATAAGGA CAATAATCGT GTCTCTTATG TGGATGGCAG CCAGTCAAGT    1920
CAGAAAAGTG AAAACTTGAC ACCAGACCAG GTTAGCCAGA AAGAAGGAAT TCAGGCTGAG    1980
CAAATTGTAA TCAAAATTAC AGATCAGGGC TATGTAACGT CACACGGTGA CCACTATCAT    2040
TACTATAATG GGAAAGTTCC TTATGATGCC CTCTTTAGTG AAGAACTCTT GATGAAGGAT    2100
CCAAACTATC AACTTAAAGA CGCTGATATT GTCAATGAAG TCAAGGGTGG TTATATCATC    2160
AAGGTCGATG GAAAATATTA TGTCTACCTG AAAGATGCAG CTCATGCTGA TAATGTTCGA    2220
ACTAAAGATG AAATCAATCG TCAAAAACAA GAACATGTCA AAGATAATGA GAAGGTTAAC    2280
TCTAATGTTG CTGTAGCAAG GTCTCAGGGA CGATATACGA CAAATGATGG TTATGTCTTT    2340
AATCCAGCTG ATATTATCGA AGATACGGGT AATGCTTATA TCGTTCCTCA TGGAGGTCAC    2400
TATCACTACA TTCCCAAAAG CGATTTATCT GCTAGTGAAT TAGCAGCAGC TAAAGCACAT    2460
CTGGCTGGAA AAAATATGCA ACCGAGTCAG TTAAGCTATT CTTCAACAGC TAGTGACAAT    2520
AACACGCAAT CTGTAGCAAA AGGATCAACT AGCAAGCCAG CAAATAAATC TGAAAATCTC    2580
CAGAGTCTTT TGAAGGAACT CTATGATTCA CCTAGCGCCC AACGTTACAG TGAATCAGAT    2640
GGCCTGGTCT TTGACCCTGC TAAGATTATC AGTCGTACAC CAAATGGAGT TGCGATTCCG    2700
CATGGCGACC ATTACCACTT TATTCCTTAC AGCAAGCTTT CTGCTTTAGA AGAAAAGATT    2760
GCCAGAATGG TGCCTATCAG TGGAACTGGT TCTACAGTTT CTACAAATGC AAAACCTAAT    2820
GAAGTAGTGT CTAGTCTAGG CAGTCTTTCA AGCAATCCTT CTTCTTTAAC GACAAGTAAG    2880
GAGCTCTCTT CAGCATCTGA TGGTTATATT TTTAATCCAA AAGATATCGT TGAAGAAACG    2940
GCTACAGCTT ATATTGTAAG ACATGGTGAT CATTTCCATT ACATTCCAAA ATCAAATCAA    3000
ATTGGGCAAC CGACTCTTCC AAACAATAGT CTAGCAACAC CTTCTCCATC TCTTCCAATC    3060
AATCCAGGAA CTTCACATGA GAAACATGAA GAAGATGGAT ACGGATTTGA TGCTAATCGT    3120
ATTATCGCTG AAGATGAATC AGGTTTTGTC ATGAGTCACG GAGACCACAA TCATTATTTC    3180
TTCAAGAAGG ACTTGACAGA AGAGCAAATT AAGGCTGCGC AAAAACATTT AGAGGAAGTT    3240
AAAACTAGTC ATAATGGATT AGATTCTTTG TCATCTCATG AACAGGATTA TCCAGGTAAT    3300
GCCAAAGAAA TGAAAGATTT AGATAAAAAA ATCGAAGAAA AAATTGCTGG CATTATGAAA    3360
```

```
CAATATGGTG TCAAACGTGA AAGTATTGTC GTGAATAAAG AAAAAAATGC GATTATTTAT     3420
CCGCATGGAG ATCACCATCA TGCAGATCCG ATTGATGAAC ATAAACCGGT TGGAATTGGT     3480
CATTCTCACA GTAACTATGA ACTGTTTAAA CCCGAAGAAG GAGTTGCTAA AAAAGAAGGG     3540
AATAAAGTTT ATACTGGAGA AGAATTAACG AATGTTGTTA ATTTGTTAAA AAATAGTACG     3600
TTTAATAATC AAAACTTTAC TCTAGCCAAT GGTCAAAAAC GCGTTTCTTT TAGTTTTCCG     3660
CCTGAATTGG AGAAAAAATT AGGTATCAAT ATGCTAGTAA AATTAATAAC ACCAGATGGA     3720
AAAGTATTGG AGAAAGTATC TGGTAAAGTA TTTGGAGAAG GAGTAGGGAA TATTGCAAAC     3780
TTTGAATTAG ATCAACCTTA TTTACCAGGA CAAACATTTA AGTATACTAT CGCTTCAAAA     3840
GATTATCCAG AAGTAAGTTA TGATGGTACA TTTACAGTTC CAACCTCTTT AGCTTACAAA     3900
ATGGCCAGTC AAACGATTTT CTATCCTTTC CATGCAGGGG ATACTTATTT AAGAGTGAAC     3960
CCTCAATTTG CAGTGCCTAA AGGAACTGAT GCTTTAGTCA GAGTGTTTGA TGAATTTCAT     4020
GGAAATGCTT ATTTAGAAAA TAACTATAAA GTTGGTGAAA TCAAATTACC GATTCCGAAA     4080
TTAAACCAAG GAACAACCAG AACGGCCGGA AATAAAATTC CTGTAACCTT CATGGCAAAT     4140
GCTTATTTGG ACAATCAATC GACTTATATT GTGGAAGTAC CTATCTTGGA AAAAGAAAAT     4200
CAAACTGATA AACCAAGTAT TCTACCACAA TTTAAAAGGA ATAAAGCACA AGAAAACTCA     4260
AAACTTGATG AAAAGGTAGA AGAACCAAAG ACTAGTGAGA AGGTAGAAAA AGAAAAACTT     4320
TCTGAAACTG GGAATAGTAC TAGTAATTCA ACGTTAGAAG AAGTTCCTAC AGTGGATCCT     4380
GTACAAGAAA AAGTAGCAAA ATTTGCTGAA AGTTATGGGA TGAAGCTAGA AAATGTCTTG     4440
TTTAATATGG ACGGAACAAT TGAATTATAT TTACCATCAG GAGAAGTCAT TAAAAAGAAT     4500
ATGGCAGATT TTACAGGAGA AGCACCTCAA GGAAATGGTG AAAATAAACC ATCTGAAAAT     4560
GGAAAAGTAT CTACTGGAAC AGTTGAGAAC CAACCAACAG AAAATAAACC AGCAGATTCT     4620
TTACCAGAGG CACCAAACGA AAAACCTGTA AAACCAGAAA ACTCAACGGA TAATGGAATG     4680
TTGAATCCAG AAGGGAATGT GGGGAGTGAC CCTATGTTAG ATCCAGCATT AGAGGAAGCT     4740
CCAGCAGTAG ATCCTGTACA AGAAAAATTA GAAAAATTTA CAGCTAGTTA CGGATTAGGC     4800
TTAGATAGTG TTATATTCAA TATGGATGGA ACGATTGAAT TAAGATTGCC AAGTGGAGAA     4860
GTGATAAAAA AGAATTTATC TGATTTCATA GCGTAAGGAA TAGCAGTAGA AAAAGTCTGA     4920
ATCAAAAATG AAGTTCTCTC AAAAGTTAGA AATAAAACTC TGACTTTGGG AGAATTTCAT     4980
TTTATTATTA ATATATAAAA TTTCTTGACA TACAACTTAA AAAGAGGTGG AATATTTACT     5040
AGTTAATT    (SEQ ID NO : 11)                                          5048
```

FIGURE 14

```
CAGAGATCTT AGTGAATCAA ATATACTTAA GAAAAGAGGA AAGAATGAAA ATCAATAAAA      60
AATATCTAGC TGGGTCAGTA GCTACACTTG TTTTAAGTGT CTGTGCTTAT GAACTAGGTT     120
TGCATCAAGC TCAAACTGTA AAAGAAAATA ATCGTGTTTC CTATATAGAT GGAAAACAAG     180
CGACGCAAAA AACGGAGAAT TTGACTCCTG ATGAGGTTAG CAAGCGTGAA GGAATCAACG     240
CCGAACAAAT CGTCATCAAG ATTACGGATC AAGGTTATGT GACCTCTCAT GGAGACCATT     300
ATCATTACTA TAATGGCAAG GTCCCTTATG ATGCCATCAT CAGTGAAGAG CTCCTCATGA     360
AAGATCCGAA TTATCAGTTG AAGGATTCAG ACATTGTCAA TGAAATCAAG GGTGGTTATG     420
TCATTAAGGT AAACGGTAAA TACTATGTTT ACCTTAAGGA TGCAGCTCAT GCGGATAATG     480
TCCGTACAAA AGAAGAAATC AATCGGCAAA AACAAGAACA TAGTCAGCAT CGTGAAGGAG     540
GGACTTCAGC AAACGATGGT GCGGTAGCCT TTGCACGTTC ACAGGGACGC TACACCACAG     600
ATGATGGTTA TATCTTCAAT GCATCTGATA TCATCGAAGA TACGGGCGAT GCCTATATCG     660
TTCCTCATGG AGATCATTAC CATTACATTC CTAAGAATGA GTTATCAGCT AGCGAGTTGG     720
CTGCTGCAGA AGCCTTCCTA TCTGGTCGGG AAAATCTGTC AAATTTAAGA ACCTATCGCC     780
GACAAAATAG CGATAACACT CCAAGAACAA ACTGGGTACC TTCTGTAAGC AATCCAGGAA     840
CTACAAATAC TAACACAAGC AACAACAGCA ACACTAACAG TCAAGCAAGT CAAAGTAATG     900
ACATTGATAG TCTCTTGAAA CAGCTCTACA AACTGCCTTT GAGTCAACGC CATGTAGAAT     960
CTGATGGCCT TATTTTCGAC CCAGCGCAAA TCACAAGTCG AACCGCCAGA GGTGTAGCTG    1020
TCCCTCATGG TAACCATTAC CACTTTATCC CTTATGAACA AATGTCTGAA TTGGAAAAAC    1080
GAATTGCTCG TATTATTCCC CTTCGTTATC GTTCAAACCA TTGGGTACCA GATTCAAGAC    1140
CAGAAGAACC AAGTCCACAA CCGACTCCAG AACCTAGTCC AAGTCCGCAA CCTGCACCAA    1200
ATCCTCAACC AGCTCCAAGC AATCCAATTG ATGAGAAATT GGTCAAAGAA GCTGTTCGAA    1260
AAGTAGGCGA TGGTTATGTC TTTGAGGAGA ATGGAGTTTC TCGTTATATC CCAGCCAAGA    1320
ATCTTTCAGC AGAAACAGCA GCAGGCATTG ATAGCAAACT GGCCAAGCAG GAAAGTTTAT    1380
CTCATAAGCT AGGAGCTAAG AAAACTGACC TCCCATCTAG TGATCGAGAA TTTTACAATA    1440
AGGCTTATGA CTTACTAGCA AGAATTCACC AAGATTTACT TGATAATAAA GGTCGACAAG    1500
TTGATTTTGA GGCTTTGGAT AACCTGTTGG AACGACTCAA GGATGTCTCA AGTGATAAAG    1560
TCAAGTTAGT GGATGATATT CTTGCCTTCT TAGCTCCGAT TCGTCATCCA GAACGTTTAG    1620
GAAAACCAAA TGCGCAAATT ACCTACACTG ATGATGAGAT TCAAGTAGCC AAGTTGGCAG    1680
GCAAGTACAC AACAGAAGAC GGTTATATCT TTGATCCTCG TGATATAACC AGTGATGAGG    1740
GGGATGCCTA TGTAACTCCA CATATGACCC ATAGCCACTG GATTAAAAAA GATAGTTTGT    1800
CTGAAGCTGA GAGAGCGGCA GCCCAGGCTT ATGCTAAAGA GAAAGGTTTG ACCCCTCCTT    1860
CGACAGACCA TCAGGATTCA GGAAATACTG AGGCAAAAGG AGCAGAAGCT ATCTACAACC    1920
GCGTGAAAGC AGCTAAGAAG GTGCCACTTG ATCGTATGCC TTACAATCTT CAATATACTG    1980
TAGAAGTCAA AAACGGTAGT TTAATCATAC CTCATTATGA CCATTACCAT AACATCAAAT    2040
TTGAGTGGTT TGACGAAGGC CTTTATGAGG CACCTAAGGG GTATACTCTT GAGGATCTTT    2100
TGGCGACTGT CAAGTACTAT GTCGAACATC CAAACGAACG TCCGCATTCA GATAATGGTT    2160
TTGGTAACGC TAGCGACCAT GTTCAAAGAA ACAAAAATGG TCAAGCTGAT ACCAATCAAA    2220
CGGAAAAACC AAGCGAGGAG AAACCTCAGA CAGAAAAACC TGAGGAAGAA ACCCCTCGAG    2280
AAGAGAAACC ACAAAGCGAG AAACCAGAGT CTCCAAAACC AACAGAGGAA CCAGAAGAAG    2340
AATCACCAGA GGAATCAGAA GAACCTCAGG TCGAGACTGA AAAGGTTGAA GAAAAACTGA    2400
GAGAGGCTGA AGATTTACTT GGAAAAATCC AGGATCCAAT TATCAAGTCC AATGCCAAAG    2460
AGACTCTCAC AGGATTAAAA AATAATTTAC TATTTGGCAC CCAGGACAAC AATACTATTA    2520
TGGCAGAAGC TGAAAAACTA TTGGCTTTAT TAAAGGAGAG TAAGTAAAGG TAGCAGCATT    2580
TTCTAACTCC TAAAAACAGG ATAGGAGAAC GGGAAAACGA AAAATGAGAG CAGAATGTGA    2640
GTTCTAG    (SED ID NO : 12)                                           2647
```

FIGURE 15

```
GGGTCTTAAA ACTCTGAATC CTTTAGAGGC AGACCCACAA AATGACAAGA CCTATTTAGA      60
AAATCTGGAA GAAAATATGA GTGTTCTAGC AGAAGAATTA AAGTGAGGAA AGAATGAAAA     120
TCAATAAAAA ATATCTAGCA GGTTCAGTGG CAGTCCTTGC CCTAAGTGTT TGTTCCTATG     180
AACTTGGTCG TCACCAAGCT GGTCAGGTTA AGAAAGAGTC TAATCGAGTT TCTTATATAG     240
ATGGTGATCA GGCTGGTCAA AAGGCAGAAA ATTTGACACC AGATGAAGTC AGTAAGAGAG     300
AGGGGATCAA CGCCGAACAA ATTGTTATCA AGATTACGGA TCAAGGTTAT GTGACCTCTC     360
ATGGAGACCA TTATCATTAC TATAATGGCA AGGTTCCTTA TGATGCCATC ATCAGTGAAG     420
AACTTCTCAT GAAAGATCCG AATTATCAGT TGAAGGATTC AGACATTGTC AATGAAATCA     480
AGGGTGGCTA TGTGATTAAG GTAGACGGAA AATACTATGT TTACCTTAAA GATGCGGCCC     540
ATGCGGACAA TATTCGGACA AAAGAAGAGA TTAAACGTCA GAAGCAGGAA CACAGTCATA     600
ATCATAACTC AAGAGCAGAT AATGCTGTTG CTGCAGCCAG AGCCCAAGGA CGTTATACAA     660
CGGATGATGG GTATATCTTC AATGCATCTG ATATCATTGA GGACACGGGT GATGCTTATA     720
TCGTTCCTCA CGGCGACCAT TACCATTACA TTCCTAAGAA TGAGTTATCA GCTAGCGAGT     780
TAGCTGCTGC AGAAGCCTAT TGGAATGGGA AGCAGGGATC TCGTCCTTCT TCAAGTTCTA     840
GTTATAATGC AAATCCAGTT CAACCAAGAT TGTCAGAGAA CCACAATCTG ACTGTCACTC     900
CAACTTATCA TCAAAATCAA GGGGAAAACA TTTCAAGCCT TTTACGTGAA TTGTATGCTA     960
AACCCTTATC AGAACGCCAT GTAGAATCTG ATGGCCTTAT TTTCGACCCA GCGCAAATCA    1020
CAAGTCGAAC CGCCAGAGGT GTAGCTGTCC CTCATGGTAA CCATTACCAC TTTATCCCTT    1080
ATGAACAAAT GTCTGAATTG GAAAAACGAA TTGCTCGTAT TATTCCCCTT CGTTATCGTT    1140
CAAACCATTG GGTACCAGAT TCAAGACCAG AACAACCAAG TCCACAATCG ACTCCGGAAC    1200
CTAGTCCAAG TCTGCAACCT GCACCAAATC CTCAACCAGC TCCAAGCAAT CCAATTGATG    1260
AGAAATTGGT CAAAGAAGCT GTTCGAAAAG TAGGCGATGG TTATGTCTTT GAGGAGAATG    1320
GAGTTTCTCG TTATATCCCA GCCAAGGATC TTTCAGCAGA AACAGCAGCA GGCATTGATA    1380
GCAAACTGGC CAAGCAGGAA AGTTTATCTC ATAAGCTAGG AGCTAAGAAA ACTGACCTCC    1440
CATCTAGTGA TCGAGAATTT TACAATAAGG CTTATGACTT ACTAGCAAGA ATTCACCAAG    1500
ATTTACTTGA TAATAAAGGT CGACAAGTTG ATTTTGAGGT TTTGGATAAC CTGTTGGAAC    1560
GACTCAAGGA TGTCTCAAGT GATAAAGTCA AGTTAGTGGA TGATATTCTT GCCTTCTTAG    1620
CTCCGATTCG TCATCCAGAA CGTTTAGGAA AACCAAATGC GCAAATTACC TACACTGATG    1680
ATGAGATTCA AGTAGCCAAG TTGGCAGGCA AGTACACAAC AGAAGACGGT TATATCTTTG    1740
ATCCTCGTGA TATAACCAGT GATGAGGGGG ATGCCTATGT AACTCCACAT ATGACCCATA    1800
GCCACTGGAT TAAAAAAGAT AGTTTGTCTG AAGCTGAGAG AGCGGCAGCC CAGGCTTATG    1860
CTAAAGAGAA AGGTTTGACC CCTCCTTCGA CAGACCACCA GGATTCAGGA AATACTGAGG    1920
CAAAAGGAGC AGAAGCTATC TACAACCGCG TGAAAGCAGC TAAGAAGGTG CCACTTGATC    1980
GTATGCCTTA CAATCTTCAA TATACTGTAG AAGTCAAAAA CGGTAGTTTA ATCATACCTC    2040
ATTATGACCA TTACCATAAC ATCAAATTTG AGTGGTTTGA CGAAGGCCTT TATGAGGCAC    2100
CTAAGGGGTA TAGTCTTGAG GATCTTTTGG CGACTGTCAA GTACTATGTC GAACATCCAA    2160
ACGAACGTCC GCATTCAGAT AATGGTTTTG GTAACGCTAG TGACCATGTT CGTAAAAATA    2220
AGGCAGACCA AGATAGTAAA CCTGATGAAG ATAAGGAACA TGATGAAGTA AGTGAGCCAA    2280
CTCACCCTGA ATCTGATGAA AAAGAGAATC ACGCTGGTTT AAATCCTTCA GCAGATAATC    2340
TTTATAAACC AAGCACTGAT ACGGAAGAGA CAGAGGAAGA AGCTGAAGAT ACCACAGATG    2400
AGGCTGAAAT TCCTCAAGTA GAGAATTCTG TTATTAACGC TAAGATAGCA GATGCGGAGG    2460
CCTTGCTAGA AAAAGTAACA GATCCTAGTA TTAGACAAAA TGCTATGGAG ACATTGACTG    2520
GTCTAAAAAG TAGTCTTCTT CTCGGAACGA AAGATAATAA CACTATTTCA GCAGAAGTAG    2580
ATAGTCTCTT GGCTTTGTTA AAAGAAAGTC AACCGGCTCC TATACAGTAG TAAAATGAA     2639
(SEQ ID NO : 13)
```

FIGURE 16

```
MKINKKYLAG SVAVLALSVC SYELGRHQAG QVKKESNRVS YIDGDQAGQK      50
AENLTPDEVS KREGINAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDAII     100
SEELLMKDPN YQLKDSDIVN EIKGGYVIKV DGKYYVYLKD AAHADNIRTK     150
EEIKRQKQEH SHNHNSRADN AVAAARAQGR YTTDDGYIFN ASDIIEDTGD     200
AYIVPHGDHY HYIPKNELSA SELAAAEAYW NGKQGSRPSS SSSYNANPVQ     250
PRLSENHNLT VTPTYHQNQG ENISSLLREL YAKPLSERHV ESDGLIFDPA     300
QITSRTARGV AVPHGNHYHF IPYEQMSELE KRIARIIPLR YRSNHWVPDS     350
RPEQPSPQST PEPSPSLQPA PNPQPAPSNP IDEKLVKEAV RKVGDGYVFE     400
ENGVSRYIPA KDLSAETAAG IDSKLAKQES LSHKLGAKKT DLPSSDREFY     450
NKAYDLLARI HQDLLDNKGR QVDFEVLDNL LERLKDVSSD KVKLVDDILA     500
FLAPIRHPER LGKPNAQITY TDDEIQVAKL AGKYTTEDGY IFDPRDITSD     550
EGDAYVTPHM THSHWIKKDS LSEAERAAAQ AYAKEKGLTP PSTDHQDSGN     600
TEAKGAEAIY NRVKAAKKVP LDRMPYNLQY TVEVKNGSLI IPHYDHYHNI     650
KFEWFDEGLY EAPKGYSLED LLATVKYYVE HPNERPHSDN GFGNASDHVR     700
KNKADQDSKP DEDKEHDEVS EPTHPESDEK ENHAGLNPSA DNLYKPSTDT     750
EETEEEAEDT TDEAEIPQVE NSVINAKIAD AEALLEKVTD PSIRQNAMET     800
LTGLKSSLLL GTKDNNTISA EVDSLLALLK ESQPAPIQ                  838
(SEQ ID NO : 14)
```

FIGURE 17

```
TGTGCCTATG CACTAAACCA GCATCGTTCG CAGGAAAATA AGGACAATAA TCGTGTCTCT      60
TATGTGGATG GCAGCCAGTC AAGTCAGAAA AGTGAAAACT TGACACCAGA CCAGGTTAGC     120
CAGAAAGAAG GAATTCAGGC TGAGCAAATT GTAATCAAAA TTACAGATCA GGGCTATGTA     180
ACGTCACACG GTGATCACTA TCATTACTAT AATGGGAAAG TTCCTTATGA TGCCCTCTTT     240
AGTGAAGAAC TCTTGATGAA GGATCCAAAC TATCAACTTA AAGACGCTGA TATTGTCAAT     300
GAAGTCAAGG GTGGTTATAT CATCAAGGTC GATGGAAAAT ATTATGTCTA CCTGAAAGAT     360
GCAGCTCATG CTGATAATGT TCGAACTAAA GATGAAATCA ATCGTCAAAA ACAAGAACAT     420
GTCAAAGATA ATGAGAAGGT TAACTCTAAT GTTGCTGTAG CAAGGTCTCA GGGACGATAT     480
ACGACAAATG ATGGTTATGT CTTTAATCCA GCTGATATTA TCGAAGATAC GGGTAATGCT     540
TATATCGTTC CTCATGGAGG TCACTATCAC TACATTCCCA AAAGCGATTT ATCTGCTAGT     600
GAATTAGCAG CAGCTAAAGC ACATCTGGCT GGAAAAAATA TGCAACCGAG TCAGTTAAGC     660
TATTCTTCAA CACCTTCTCC ATCTCTTCCA ATCAATCCAG GAACTTCACA TGAGAAACAT     720
GAAGAAGATG GATACGGATT TGATGCTAAT CGTATTATCG CTGAAGATGA ATCAGGTTTT     780
GTCATGAGTC ACGGAGACCA CAATCATTAT TTCTTCAAGA AGGACTTGAC AGAAGAGCAA     840
ATTAAGGCTG CGCAAAAACA TTTAGAGGAA GTTAAAACTA GTCATAATGG ATTAGATTCT     900
TTGTCATCTC ATGAACAGGA TTATCCAAGT AATGCCAAAG AAATGAAAGA TTTAGATAAA     960
AAAATCGAAG AAAAAATTGC TGGCATTATG AAACAATATG GTGTCAAACG TGAAAGTATT    1020
GTCGTGAATA AAGAAAAAAA TGCGATTATT TATCCGCATG GAGATCACCA TCATGCAGAT    1080
CCGATTGATG AACATAAACC GGTTGGAATT GGTCATTCTC ACAGTAACTA TGAACTGTTT    1140
AAACCCGAAG AAGGAGTTGC TAAAAAAGAA GGGAATAAAG TTTATACTGG AGAAGAATTA    1200
ACGAATGTTG TTAATTTGTT AAAAAATAGT ACGTTTAATA ATCAAAACTT TACTCTAGCC    1260
AATGGTCAAA AACGCGTTTC TTTTAGTTTT CCGCCTGAAT TGGAGAAAAA ATTAGGTATC    1320
AATATGCTAG TAAAATTAAT AACACCAGAT GGAAAAGTAT TGGAGAAAGT ATCTGGTAAA    1380
GTATTTGGAG AAGGAGTAGG GAATATTGCA AACTTTGAAT TAGATCAACC TTATTTACCA    1440
GGACAAACAT TTAAGTATAC TATCGCTTCA AAAGATTATC CAGAAGTAAG TTATGATGGT    1500
ACATTTACAG TTCCAACCTC TTTAGCTTAC AAAATGGCCA GTCAAACGAT TTTCTATCCT    1560
TTCCATGCAG GGGATACTTA TTTAAGAGTG AACCCTCAAT TTGCAGTGCC TAAAGGAACT    1620
GATGCTTTAG TCAGAGTGTT TGATGAATTT CATGGAAATG CTTATTTAGA AAATAACTAT    1680
AAAGTTGGTG AAATCAAATT ACCGATTCCG AAATTAAACC AAGGAACAAC CAGAACGGCC    1740
GGAAATAAAA TTCCTGTAAC CTTCATGGCA AATGCTTATT TGGACAATCA ATCGACTTAT    1800
ATTGTGGAAG TACCTATCTT GGAAAAAGAA AATCAAACTG ATAAACCAAG TATTCTACCA    1860
CAATTTAAAA GGAATAAAGC ACAAGAAAAC TCAAAACTTG ATGAAAAGGT AGAAGAACCA    1920
AAGACTAGTG AGAAGGTAGA AAAAGAAAAA CTTTCTGAAA CTGGGAATAG TACTAGTAAT    1980
TCAACGTTAG AAGAAGTTCC TACAGTGGAT CCTGTACAAG AAAAAGTAGC AAAATTTGCT    2040
GAAAGTTATG GGATGAAGCT AGAAAATGTC TTGTTTAATA TGGACGGAAC AATTGAATTA    2100
TATTTACCAT CGGGAGAAGT CATTAAAAAG AATATGGCAG ATTTTACAGG AGAAGCACCT    2160
CAAGGAAATG GTGAAAATAA ACCATCTGAA AATGGAAAAG TATCTACTGG AACAGTTGAG    2220
AACCAACCAA CAGAAAATAA ACCAGCAGAT TCTTTACCAG AGGCACCAAA CGAAAAACCT    2280
GTAAAACCAG AAAACTCAAC GGATAATGGA ATGTTGAATC CAGAAGGGAA TGTGGGGAGT    2340
GACCCTATGT TAGATTCAGC ATTAGAGGAA GCTCCAGCAG TAGATCCTGT ACAAGAAAAA    2400
TTAGAAAAAT TTACAGCTAG TTACGGATTA GGCTTAGATA GTGTTATATT CAATATGGAT    2460
GGAACGATTG AATTAAGATT GCCAAGTGGA GAAGTGATAA AAAAGAATTT ATTGATCTCA    2520
TAGCGTAA    (SEQ ID NO : 15)                                          2528
```

FIGURE 18

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN     150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS     200
ELAAAKAHLA GKNMQPSQLS YSSTPSPSLP INPGTSHEKH EEDGYGFDAN     250
RIIAEDESGF VMSHGDHNHY FFKKDLTEEQ IKAAQKHLEE VKTSHNGLDS     300
LSSHEQDYPS NAKEMKDLDK KIEEKIAGIM KQYGVKRESI VVNKEKNAII     350
YPHGDHHHAD PIDEHKPVGI GHSHSNYELF KPEEGVAKKE GNKVYTGEEL     400
TNVVNLLKNS TFNNQNFTLA NGQKRVSFSF PPELEKKLGI NMLVKLITPD     450
GKVLEKVSGK VFGEGVGNIA NFELDQPYLP GQTFKYTIAS KDYPEVSYDG     500
TFTVPTSLAY KMASQTIFYP FHAGDTYLRV NPQFAVPKGT DALVRVFDEF     550
HGNAYLENNY KVGEIKLPIP KLNQGTTRTA GNKIPVTFMA NAYLDNQSTY     600
IVEVPILEKE NQTDKPSILP QFKRNKAQEN SKLDEKVEEP KTSEKVEKEK     650
LSETGNSTSN STLEEVPTVD PVQEKVAKFA ESYGMKLENV LFNMDGTIEL     700
YLPSGEVIKK NMADFTGEAP QGNGENKPSE NGKVSTGTVE NQPTENKPAD     750
SLPEAPNEKP VKPENSTDNG MLNPEGNVGS DPMLDSALEE APAVDPVQEK     800
LEKFTASYGL GLDSVIFNMD GTIELRLPSG EVIKKNLLIS                840
(SEQ ID NO : 16)
```

FIGURE 19

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN     150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS     200
ELAAAKAHLA GKNMQPSQLS YSSTASDNNT QSVAKGSTSK PANKSENLQS     250
LLKELYDSPS AQRYSESDGL VFDPAKIISR TPNGVAIPHG DHYHFIPYSK     300
LSALEEKIAR MVPISGTGST VSTNAKPNEV VSSLGSLSSN PSSLTTSKEL     350
SSASDGYIFN PKDIVEETAT AYIVRHGDHF HYIPKSNQIG QPTLPNNSLA     400
TPSPSLPINP GTSHEKHEED GYGFDANRII AEDESGFVMS HGDHNHYFFK     450
KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS HEQDYPGNAK EMKDLDKKIE     500
EKIAGIMKQY GVKRESIVVN KEKNAIIYPH GDHHHADPID EHKPVGIGHS     550
HSNYELFKPE EGVAKKEGNK VYTGEELTNV VNLLKNSTFN NQNFTLANGQ     600
KRVSFSFPPE LEKKLGINML VKLITPDGKV LEKVSGKVFG EGVGNIANFE     650
LDQPYLPGQT FKYTIASKDY PEVSYDGTFT VPTSLAYKMA SQTIFYPFHA     700
GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN AYLENNYKVG EIKLPIPKLN     750
QGTTRTAGNK IPVTFMANAY LDNQSTYIVE VPILEKENQT DKPSILPQFK     800
RNKAQENSKL DEKVEEPKTS EKVEKEKLSE TGNSTSNSTL EEVPTVDPVQ     850
EKVAKFAESY GMKLENVLFN MDGTIELYLP SGEVIKKNMA DFTGEAPQGN     900
GENKPSENGK VSTGTVENQP TENKPADSLP EAPNEKPVKP ENSTDNGMLN     950
PEGNVGSDPM LDPALEEAPA VDPVQEKLEK FTASYGLGLD SVIFNMDGTI    1000
ELRLPSGEVI KKNLSDFIA   (SEQ ID NO : 55)                   1019
```

FIGURE 20

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN     150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS     200
ELAAAKAHLA GKNMQPSQLS YSSTASDNNT QSVAKGSTSK PANKSENLQS     250
LLKELYDSPS AQRYSESDGL VFDPAKIISR TPNGVAIPHG DHYHFIPYSK     300
LSALEEKIAR MVPISGTGST VSTNAKPNEV VSSLGSLSSN PSSLTTSKEL     350
SSASDGYIFN PKDIVEETAT AYIVRHGDHF HYIPKSNQIG QPTLPNNSLA     400
TPSPSLPINP GTSHEKHEED GYGFDANRII AEDESGFVMS HGDHNHYFFK     450
KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS HEQDYPGNA                 489
(SEQ ID NO : 56)
```

FIGURE 21

```
MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS      60
QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     120
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN VAVARSQGRY     180
TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS     240
YSSTASDNNT QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR     300
TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV VSSLGSLSSN     360
PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF HYIPKSNQIG QPTLPNNSLA     420
TPSPSLPINP GTSHEKHEED GYGFDANRII AEDESGFVMS HGDHNHYFFK KDLTEEQIKA     480
AQKHLEEVKT SHNGLDSLSS HEQDYPGNA   (SEQ ID NO : 57)                    509
```

FIGURE 22

```
DLTEEQIKAA QKHLEEVKTS HNGLDSLSSH EQDYPGNAKE MKDLDKKIEE      50
KIAGIMKQYG VKRESIVVNK EKNAIIYPHG DHHHADPIDE HKPVGIGHSH     100
SNYELFKPEE GVAKKEGNKV YTGEELTNVV NLLKNSTFNN QNFTLANGQK     150
RVSFSFPPEL EKKLGINMLV KLITPDGKVL EKVSGKVFGE GVGNIANFEL     200
DQPYLPGQTF KYTIASKDYP EVSYDGTFTV PTSLAYKMAS QTIFYPFHAG     250
DTYLRVNPQF AVPKGTDALV RVFDEFHGNA YLENNYKVGE IKLPIPKLNQ     300
GTTRTAGNKI PVTFMANAYL DNQSTYIVEV PILEKENQTD KPSILPQFKR     350
NKAQENSKLD EKVEEPKTSE KVEKEKLSET GNSTSNSTLE EVPTVDPVQE     400
KVAKFAESYG MKLENVLFNM DGTIELYLPS GEVIKKNMAD FTGEAPQGNG     450
ENKPSENGKV STGTVENQPT ENKPADSLPE APNEKPVKPE NSTDNGMLNP     500
EGNVGSDPML DPALEEAPAV DPVQEKLEKF TASYGLGLDS VIFNMDGTIE     550
LRLPSGEVIK KNLSDFIAKL RYRSNHWVPD SRPEEPSPQP TPEPSPSPQP     600
APNPQPAPSN PIDEKLVKEA VRKVGDGYVF EENGVSRYIP AKNLSAETAA     650
GIDSKLAKQE SLSHKLGAKK TDLPSSDREF YNKAYDLLAR IHQDLLDNKG     700
RQVDFEALDN LLERLKDVSS DKVKLVDDIL AFLAPIRHPE RLGKPNAQIT     750
YTDDEIQVAK LAGKYTTEDG YIFDPRDITS DEGDAYVTPH MTHSHWIKKD     800
SLSEAERAAA QAYAKEKGLT PPSTDHQDSG NTEAKGAEAI YNRVKAAKKV     850
PLDRMPYNLQ YTVEVKNGSL IIPHYDHYHN IKFEWFDEGL YEAPKGYTLE     900
DLLATVKYYV EHPNERPHSD NGFGNASDHV QRNKNGQADT NQTEKPSEEK     950
PQTEKPEEET PREEKPQSEK PESPKPTEEP EEESPEESEE PQVETEKVEE    1000
KLREAEDLLG KIQDPIIKSN AKETLTGLKN NLLFGTQDNN TIMAEAEKLL    1050
ALLKESK   (SEQ ID NO : 58)                                 1057
```

FIGURE 23

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN     150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS     200
ELAAA    (SEQ ID NO :  59)                                  205
```

FIGURE 24

```
CAYELGLHQA QTVKENNRVS YIDGKQATQK TENLTPDEVS KREGINAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDAII SEELLMKDPN YQLKDSDIVN     100
EIKGGYVIKV NGKYYVYLKD AAHADNVRTK EEINRQKQEH SQHREGGTSA     150
NDGAVAFARS QGRYTTDDGY IFNASDIIED TGDAYIVPHG DHYHYIPKNE     200
LSASELAAAE AFLSGRENLS NLRTYRRQNS DNTPRTNWVP SVSNPGTTNT     250
NTSNNSNTNS QASQSNDIDS LLKQLYKLPL SQRHVESDGL IFDPAQITSR     300
TARGVAVPHG NHYHFIPYEQ MSELEKRIAR IIPLRYRSNH WVPDSRPEEP     350
SPQPTPEPSP SPQPAPNPQP APSNPIDEKL VKEAVRKVGD GYVFEENGVS     400
RYIPAKNLSA ETAAGIDSKL AKQESLSHKL GAKKTDLPSS DREFYNKAYD     450
LLARIHQDLL DNKGRQVDFE ALDNLLERLK DVSSDKVKLV DDILAFLAPI     500
RHPERLGKPN AQITYTDDEI QVAKLAGKYT TEDGYIFDPR DITSDEGDAY     550
VTPHMTHSHW IKKDSLSEAE RAAAQAYAKE KGLTPPSTDH QDSGNTEAKG     600
AEAIYNRVKA AKKVPLDRMP YNLQYTVEVK NGSLIIPHYD HYHNIKFEWF     650
DEGLYEAPKG YTLEDLLATV KYYVEHPNER PHSDNGFGNA SDHVQRNKNG     700
QADTNQTEKP SEEKPQTEKP EEETPREEKP QSEKPESPKP TEEPEEESPE     750
ESEEPQVETE KVEEKLREAE DLLGKIQDPI IKSNAKETLT GLKNNLLFGT     800
QDNNTIMAEA EKLLALLKES K    ((SEQ ID NO : 60)               821
```

FIGURE 25

```
CAYELGLHQA QTVKENNRVS YIDGKQATQK TENLTPDEVS KREGINAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDAII SEELLMKDPN YQLKDSDIVN     100
EIKGGYVIKV NGKYYVYLKD AAHADNVRTK EEINRQKQEH SQHREGGTSA     150
NDGAVAFARS QGRYTTDDGY IFNASDIIED TGDAYIVPHG DHYHYIPKNE     200
LSASELAAAE AFLSGRENLS NLRTYRRQNS DNTPRTNWVP SVSNPGTTNT     250
NTSNNSNTNS QASQSNDIDS LLKQLYKLPL SQRHVESDGL IFDPAQITSR     300
TARGVAVPHG NHYHFIPYEQ MSELEKRIAR IIPL                      334
(SEQ ID NO : 61)
```

FIGURE 26

```
RYRSNHWVPD SRPEEPSPQP TPEPSPSPQP APNPQPAPSN PIDEKLVKEA      50
VRKVGDGYVF EENGVSRYIP AKNLSAETAA GIDSKLAKQE SLSHKLGAKK     100
TDLPSSDREF YNKAYDLLAR IHQDLLDNKG RQVDFEALDN LLERLKDVSS     150
DKVKLVDDIL AFLAPIRHPE RLGKPNAQIT YTDDEIQVAK LAGKYTTEDG     200
YIFDPRDITS DEGDAYVTPH MTHSHWIKKD SLSEAERAAA QAYAKEKGLT     250
PPSTDHQDSG NTEAKGAEAI YNRVKAAKKV PLDRMPYNLQ YTVEVKNGSL     300
IIPHYDHYHN IKFEWFDEGL YEAPKGYTLE DLLATVKYYV EHPNERPHSD     350
NGFGNASDHV QRNKNGQADT NQTEKPSEEK PQTEKPEEET PREEKPQSEK     400
PESPKPTEEP EEESPEESEE PQVETEKVEE KLREAEDLLG KIQDPIIKSN     450
AKETLTGLKN NLLFGTQDNN TIMAEAEKLL ALLKESK                   487
(SEQ ID NO : 62)
```

FIGURE 27

```
AEAFLSGREN LSNLRTYRRQ NSDNTPRTNW VPSVSNPGTT NTNTSNNSNT      50
NSQASQSNDI DSLLKQLYKL PLSQRHVESD GLIFDPAQIT SRTARGVAVP     100
HGNHYHFIPY EQMSELEKRI ARIIPLRYRS NHWVPDSRPE EPSPQPTPEP     150
SPSPQPAPNP QPAPSNPIDE KLVKEAVRKV GDGYVFEENG VSRYIPAKNL     200
SAETAAGIDS KLAKQESLSH KLGAKKTDLP SSDREFYNKA YDLLARIHQD     250
LLDNKGRQVD FEALDNLLER LKDVSSDKVK LVDDILAFLA PIRHPERLGK     300
PNAQITYTDD EIQVAKLAGK YTTEDGYIFD PRDITSDEGD AYVTPHMTHS     350
HWIKKDSLSE AERAAAQAYA KEKGLTPPST DHQDSGNTEA KGAEAIYNRV     400
KAAKKVPLDR MPYNLQYTVE VKNGSLIIPH YDHYHNIKFE WFDEGLYEAP     450
KGYTLEDLLA TVKYYVEHPN ERPHSDNGFG NASDHVQRNK NGQADTNQTE     500
KPSEEKPQTE KPEEETPREE KPQSEKPESP KPTEEPEEES PEESEEPQVE     550
TEKVEEKLRE AEDLLGKIQD PIIKSNAKET LTGLKNNLLF GTQDNNTIMA     600
EAEKLLALLK ESK    (SEQ ID NO : 63)                         613
```

FIGURE 28

```
DLTEEQIKAA QKHLEEVKTS HNGLDSLSSH EQDYPGNAKE MKDLDKKIEE      50
KIAGIMKQYG VKRESIVVNK EKNAIIYPHG DHHHADPIDE HKPVGIGHSH     100
SNYELFKPEE GVAKKEGNKV YTGEELTNVV NLLKNSTFNN QNFTLANGQK     150
RVSFSFPPEL EKKLGINMLV KLITPDGKVL EKVSGKVFGE GVGNIANFEL     200
DQPYLPGQTF KYTIASKDYP EVSYDGTFTV PTSLAYKMAS QTIFYPFHAG     250
DTYLRVNPQF AVPKGTDALV RVFDEFHGNA YLENNYKVGE IKLPIPKLNQ     300
GTTRTAGNKI PVTFMANAYL DNQSTYIVEV PILEKENQTD KPSILPQFKR     350
NKAQENSKLD EKVEEPKTSE KVEKEKLSET GNSTSNSTLE EVPTVDPVQE     400
KVAKFAESYG MKLENVLFNM DGTIELYLPS GEVIKKNMAD FTGEAPQGNG     450
ENKPSENGKV STGTVENQPT ENKPADSLPE APNEKPVKPE NSTDNGMLNP     500
EGNVGSDPML DPALEEAPAV DPVQEKLEKF TASYGLGLDS VIFNMDGTIE     550
LRLPSGEVIK KNLSDFIA    (SEQ ID NO : 64)                    568
```

FIGURE 29

```
DLTEEQIKAA QKHLEEVKTS HNGLDSLSSH EQDYPGNAKE MKDLDKKIEE      50
KIAGIMKQYG VKRESIVVNK EKNAIIYPHG DHHHADPIDE HKPVGIGHSH     100
SNYELFKPEE GVAKKEGNKV YTGEELTNVV NLLKNSTFNN QNFTLANGQK     150
RVSFSFPPEL EKKLGINMLV KLITPDGKVL EKVSGKVFGE GVGNIANFEL     200
DQPYLPGQTF KYTIASKDYP EVSYDGTFTV PTSLAYKMAS QTIFYPFHAG     250
DTYLRVNPQF AVPKGTDALV RVFDEFHGNA YLENNYKVGE IKLPIPKLNQ     300
GTTRTAGNKI PVTFMANAYL DNQSTYIVE    (SEQ ID NO : 65)       329
```

FIGURE 30

```
EVPILEKENQ TDKPSILPQF KRNKAQENSK LDEKVEEPKT SEKVEKEKLS      50
ETGNSTSNST LEEVPTVDPV QEKVAKFAES YGMKLENVLF NMDGTIELYL     100
PSGEVIKKNM ADFTGEAPQG NGENKPSENG KVSTGTVENQ PTENKPADSL     150
PEAPNEKPVK PENSTDNGML NPEGNVGSDP MLDPALEEAP AVDPVQEKLE     200
KFTASYGLGL DSVIFNMDGT IELRLPSGEV IKKNLSDFIA                240
(SEQ ID NO : 66)
```

FIGURE 31

```
DIDSLLKQLY KLPLSQRHVE SDGLIFDPAQ ITSRTARGVA VPHGNHYHFI      50
PYEQMSELEK RIARIIPLRY RSNHWVPDSR PEEPSPQPTP EPSPSPQPAP     100
NPQPAPSNPI DEKLVKEAVR KVGDGYVFEE NGVSRYIPAK NLSAETAAGI     150
DSKLAKQESL SHKLGAKKTD LPSSDREFYN KAYDLLARIH QDLLDNKGRQ     200
VDFEALDNLL ERLKDVSSDK VKLVDDILAF LAPIRHPERL GKPNAQITYT     250
DDEIQVAKLA GKYTTEDGYI FDPRDITSDE GDAYVTPHMT HSHWIKKDSL     300
SEAERAAAQA YAKEKGLTPP STDHQDSGNT EAKGAEAIYN RVKAAKKVPL     350
DRMPYNLQYT VEVKNGSLII PHYDHYHNIK FEWFDEGLYE APKGYTLEDL     400
LATVKYYVEH PNERPHSDNG FGNASDHVQR NKNGQADTNQ TEKPSEEKPQ     450
TEKPEEETPR EEKPQSEKPE SPKPTEEPEE ESPEESEEPQ VETEKVEEKL     500
REAEDLLGKI QDPIIKSNAK ETLTGLKNNL LFGTQDNNTI MAEAEKLLAL     550
LKESK    (SEQ ID NO : 67)                                  555
```

FIGURE 32

```
DIDSLLKQLY KLPLSQRHVE SDGLIFDPAQ ITSRTARGVA VPHGNHYHFI      50
PYEQMSELEK RIARIIPLRY RSNHWVPDSR PEEPSPQPTP EPSPSPQPAP     100
NPQPAPSNPI DEKLVKEAVR KVGDGYVFEE NGVSRYIPAK NLSAETAAGI     150
DSKLAKQESL SHKLGAKKTD LPSSDREFYN KAYDLLARIH QDLLDNKGRQ     200
VDFEALDNLL ERLKDVSSDK VKLVDDILAF LAPIRHPERL GKPNAQITYT     250
DDEIQVAKLA GKYTTEDGYI FDPRDITSDE GDAYVTPHMT HSHWIKKDSL     300
SEAERAAAQA YAKEKGLTPP STDHQDSGNT EAKGAEAIYN RVKAAKKVPL     350
DRMPYNLQYT VEVKNGSLII PHYDHYHNIK FEWFDEGLYE APKGYTLEDL     400
LATVKYYVEH PNERPHSDNG FGNASDHV    (SEQ ID NO : 68)        428
```

FIGURE 33

```
GLYEAPKGYT LEDLLATVKY YVEHPNERPH SDNGFGNASD HVQRNKNGQA      50
DTNQTEKPSE EKPQTEKPEE ETPREEKPQS EKPESPKPTE EPEEESPEES     100
EEPQVETEKV EEKLREAEDL L    (SEQ ID NO : 69)                121
```

FIGURE 34

```
ASDHVQRNKN GQADTNQTEK PSEEKPQTEK PEEETPREEK PQSEKPESPK      50
PTEEPEEESP EESEEPQVET EKVEEKLREA EDLLGKIQDP IIKSNAKETL     100
TGLKNNLLFG TQDNNTIMAE AEKLLALLKE SK                        132
(SEQ ID NO : 70)
```

FIGURE 35

```
DIDSLLKQLY KLPLSQRHVE SDGLIFDPAQ ITSRTARGVA VPHGNHYHFI      50
PYEQMSELEK RIARIIPLRY RSNHWVPDSR PEEPSPQPTP EPSPSPQPAP     100
NPQPAPSNPI DEKLVKEAVR KVGDGYVFEE NGVSRYIPAK NLSAETAAGI     150
DSKLAKQESL SHKLGAKKTD LPSSDREFYN KAYDLLARIH QDLLDNKGRQ     200
VDFEALDNLL ERLKDVSSDK VKLVDD    (SEQ ID NO : 71)           226
```

FIGURE 36

```
DILAFLAPIR HPERLGKPNA QITYTDDEIQ VAKLAGKYTT EDGYIFDPRD      50
ITSDEGDAYV TPHMTHSHWI KKDSLSEAER AAAQAYAKEK GLTPPSTDHQ     100
DSGNTEAKGA EAIYNRVKAA KKVPLDRMPY NLQYTVEVKN GSLIIPHYDH     150
YHNIKFEWFD EGLYEAPKGY TLEDLLATVK YYVEHPNERP HSDNGFGNAS     200
DHV    (SEQ ID NO : 72)                                    203
```

FIGURE 37

```
CSYELGRHQA GQVKKESNRV SYIDGDQAGQ KAENLTPDEV SKREGINAEQ      50
IVIKITDQGY VTSHGDHYHY YNGKVPYDAI ISEELLMKDP NYQLKDSDIV     100
NEIKGGYVIK VDGKYYVYLK DAAHADNIRT KEEIKRQKQE HSHNHNSRAD     150
NAVAAARAQG RYTTDDGYIF NASDIIEDTG DAYIVPHGDH YHYIPKNELS     200
ASELAAAEAY WNGKQGSRPS SSSSYNANPV QPRLSENHNL TVTPTYHQNQ     250
GENISSLLRE LYAKPLSERH VESDGLIFDP AQITSRTARG VAVPHGNHYH     300
FIPYEQMSEL EKRIARIIPL RYRSNHWVPD SRPEQPSPQS TPEPSPSLQP     350
APNPQPAPSN PIDEKLVKEA VRKVGDGYVF EENGVSRYIP AKDLSAETAA     400
GIDSKLAKQE SLSHKLGAKK TDLPSSDREF YNKAYDLLAR IHQDLLDNKG     450
RQVDFEVLDN LLERLKDVSS DKVKLVDDIL AFLAPIRHPE RLGKPNAQIT     500
YTDDEIQVAK LAGKYTTEDG YIFDPRDITS DEGDAYVTPH MTHSHWIKKD     550
SLSEAERAAA QAYAKEKGLT PPSTDHQDSG NTEAKGAEAI YNRVKAAKKV     600
PLDRMPYNLQ YTVEVKNGSL IIPHYDHYHN IKFEWFDEGL YEAPKGYSLE     650
DLLATVKYYV EHPNERPHSD NGFGNASDHV RKNKADQDSK PDEDKEHDEV     700
SEPTHPESDE KENHAGLNPS ADNLYKPSTD TEETEEEAED TTDEAEIPQV     750
ENSVINAKIA DAEALLEKVT DPSIRQNAME TLTGLKSSLL LGTKDNNTIS     800
AEVDSLLALL KESQPAPIQ    (SEQ ID NO : 73)                   819
```

FIGURE 38

```
ENISSLLREL YAKPLSERHV ESDGLIFDPA QITSRTARGV AVPHGNHYHF      50
IPYEQMSELE KRIARIIPLR YRSNHWVPDS RPEQPSPQST PEPSPSLQPA     100
PNPQPAPSNP IDEKLVKEAV RKVGDGYVFE ENGVSRYIPA KDLSAETAAG     150
IDSKLAKQES LSHKLGAKKT DLPSSDREFY NKAYDLLARI HQDLLDNKGR     200
QVDFEVLDNL LERLKDVSSD KVKLVDDILA FLAPIRHPER LGKPNAQITY     250
TDDEIQVAKL AGKYTTEDGY IFDPRDITSD EGDAYVTPHM THSHWIKKDS     300
LSEAERAAAQ AYAKEKGLTP PSTDHQDSGN TEAKGAEAIY NRVKAAKKVP     350
LDRMPYNLQY TVEVKNGSLI IPHYDHYHNI KFEWFDEGLY EAPKGYSLED     400
LLATVKYYVE HPNERPHSDN GFGNASDHVR KNKADQDSKP DEDKEHDEVS     450
EPTHPESDEK ENHAGLNPSA DNLYKPSTDT EETEEEAEDT TDEAEIPQVE     500
NSVINAKIAD AEALLEKVTD PSIRQNAMET LTGLKSSLLL GTKDNNTISA     550
EVDSLLALLK ESQPAPIQ   (SEQ ID NO : 74)                     568
```

FIGURE 39

```
VRKNKADQDS KPDEDKEHDE VSEPTHPESD EKENHAGLNP SADNLYKPST        50
DTEETEEEAE DTTDEAEIPQ VENSVINAKI ADAEALLEKV TDPSIRQNAM       100
ETLTGLKSSL LLGTKDNNTI SAEVDSLLAL LKESQPAPIQ                  140
(SEQ ID NO : 75)
```

FIGURE 40

```
GACTTGACAG AAGAGCAAAT TAAGGCTGCG CAAAAACATT TAGAGGAAGT        50
TAAAACTAGT CATAATGGAT TAGATTCTTT GTCATCTCAT GAACAGGATT       100
ATCCAGGTAA TGCCAAAGAA ATGAAAGATT TAGATAAAAA AATCGAAGAA       150
AAAATTGCTG GCATTATGAA ACAATATGGT GTCAAACGTG AAAGTATTGT       200
CGTGAATAAA GAAAAAAATG CGATTATTTA TCCGCATGGA GATCACCATC       250
ATGCAGATCC GATTGATGAA CATAAACCGG TTGGAATTGG TCATTCTCAC       300
AGTAACTATG AACTGTTTAA ACCCGAAGAA GGAGTTGCTA AAAAAGAAGG       350
GAATAAAGTT TATACTGGAG AAGAATTAAC GAATGTTGTT AATTTGTTAA       400
AAAATAGTAC GTTTAATAAT CAAAACTTTA CTCTAGCCAA TGGTCAAAAA       450
CGCGTTTCTT TTAGTTTTCC GCCTGAATTG GAGAAAAAAT TAGGTATCAA       500
TATGCTAGTA AAATTAATAA CACCAGATGG AAAAGTATTG GAGAAAGTAT       550
CTGGTAAAGT ATTTGGAGAA GGAGTAGGGA ATATTGCAAA CTTTGAATTA       600
GATCAACCTT ATTTACCAGG ACAAACATTT AAGTATACTA TCGCTTCAAA       650
AGATTATCCA GAAGTAAGTT ATGATGGTAC ATTTACAGTT CCAACCTCTT       700
TAGCTTACAA AATGGCCAGT CAAACGATTT TCTATCCTTT CCATGCAGGG       750
GATACTTATT TAAGAGTGAA CCCTCAATTT GCAGTGCCTA AAGGAACTGA       800
TGCTTTAGTC AGAGTGTTTG ATGAATTTCA TGGAAATGCT TATTTAGAAA       850
ATAACTATAA AGTTGGTGAA ATCAAATTAC CGATTCCGAA ATTAAACCAA       900
GGAACAACCA GAACGGCCGG AAATAAAATT CCTGTAACCT TCATGGCAAA       950
TGCTTATTTG GACAATCAAT CGACTTATAT TGTGGAAGTA CCTATCTTGG      1000
AAAAAGAAAA TCAAACTGAT AAACCAAGTA TTCTACCACA ATTTAAAAGG      1050
AATAAAGCAC AAGAAAACTC AAAACTTGAT GAAAAGGTAG AAGAACCAAA      1100
GACTAGTGAG AAGGTAGAAA AAGAAAAACT TTCTGAAACT GGGAATAGTA      1150
CTAGTAATTC AACGTTAGAA GAAGTTCCTA CAGTGGATCC TGTACAAGAA      1200
AAAGTAGCAA AATTTGCTGA AAGTTATGGG ATGAAGCTAG AAAATGTCTT      1250
GTTTAATATG GACGGAACAA TTGAATTATA TTTACCATCA GGAGAAGTCA      1300
TTAAAAAGAA TATGGCAGAT TTTACAGGAG AAGCACCTCA AGGAAATGGT      1350
GAAAATAAAC CATCTGAAAA TGGAAAAGTA TCTACTGGAA CAGTTGAGAA      1400
CCAACCAACA GAAAATAAAC CAGCAGATTC TTTACCAGAG GCACCAAACG      1450
AAAAACCTGT AAAACCAGAA AACTCAACGG ATAATGGAAT GTTGAATCCA      1500
GAAGGGAATG TGGGGAGTGA CCCTATGTTA GATCCAGCAT TAGAGGAAGC      1550
TCCAGCAGTA GATCCTGTAC AAGAAAAATT AGAAAAATTT ACAGCTAGTT      1600
ACGGATTAGG CTTAGATAGT GTTATATTCA ATATGGATGG AACGATTGAA      1650
TTAAGATTGC CAAGTGGAGA AGTGATAAAA AAGAATTTAT CTGATTTCAT      1700
AGCGAAGCTT CGTTATCGTT CAAACCATTG GGTACCAGAT TCAAGACCAG      1750
AAGAACCAAG TCCACAACCG ACTCCAGAAC CTAGTCCAAG TCCGCAACCT      1800
GCACCAAATC CTCAACCAGC TCCAAGCAAT CCAATTGATG AGAAATTGGT      1850
CAAAGAAGCT GTTCGAAAAG TAGGCGATGG TTATGTCTTT GAGGAGAATG      1900
GAGTTTCTCG TTATATCCCA GCCAAGAATC TTTCAGCAGA AACAGCAGCA      1950
GGCATTGATA GCAAACTGGC CAAGCAGGAA AGTTTATCTC ATAAGCTAGG      2000
AGCTAAGAAA ACTGACCTCC CATCTAGTGA TCGAGAATTT TACAATAAGG      2050
CTTATGACTT ACTAGCAAGA ATTCACCAAG ATTTACTTGA TAATAAAGGT      2100
CGACAAGTTG ATTTTGAGGC TTTGGATAAC CTGTTGGAAC GACTCAAGGA      2150
TGTCTCAAGT GATAAAGTCA AGTTAGTGGA TGATATTCTT GCCTTCTTAG      2200
CTCCGATTCG TCATCCAGAA CGTTTAGGAA AACCAAATGC GCAAATTACC      2250
TACACTGATG ATGAGATTCA AGTAGCCAAG TTGGCAGGCA AGTACACAAC      2300
AGAAGACGGT TATATCTTTG ATCCTCGTGA TATAACCAGT GATGAGGGGG      2350
ATGCCTATGT AACTCCACAT ATGACCCATA GCCACTGGAT TAAAAAAGAT      2400
```

```
AGTTTGTCTG AAGCTGAGAG AGCGGCAGCC CAGGCTTATG CTAAAGAGAA     2450
AGGTTTGACC CCTCCTTCGA CAGACCATCA GGATTCAGGA AATACTGAGG     2500
CAAAAGGAGC AGAAGCTATC TACAACCGCG TGAAAGCAGC TAAGAAGGTG     2550
CCACTTGATC GTATGCCTTA CAATCTTCAA TATACTGTAG AAGTCAAAAA     2600
CGGTAGTTTA ATCATACCTC ATTATGACCA TTACCATAAC ATCAAATTTG     2650
AGTGGTTTGA CGAAGGCCTT TATGAGGCAC CTAAGGGGTA TACTCTTGAG     2700
GATCTTTTGG CGACTGTCAA GTACTATGTC GAACATCCAA ACGAACGTCC     2750
GCATTCAGAT AATGGTTTTG GTAACGCTAG CGACCATGTT CAAAGAAACA     2800
AAAATGGTCA AGCTGATACC AATCAAACGG AAAAACCAAG CGAGGAGAAA     2850
CCTCAGACAG AAAAACCTGA GGAAGAAACC CCTCGAGAAG AGAAACCACA     2900
AAGCGAGAAA CCAGAGTCTC CAAAACCAAC AGAGGAACCA GAAGAAGAAT     2950
CACCAGAGGA ATCAGAAGAA CCTCAGGTCG AGACTGAAAA GGTTGAAGAA     3000
AAACTGAGAG AGGCTGAAGA TTTACTTGGA AAAATCCAGG ATCCAATTAT     3050
CAAGTCCAAT GCCAAAGAGA CTCTCACAGG ATTAAAAAAT AATTTACTAT     3100
TTGGCACCCA GGACAACAAT ACTATTATGG CAGAAGCTGA AAAACTATTG     3150
GCTTTATTAA AGGAGAGTAA G    (SEQ ID NO : 76)                3171
```

FIGURE 41

```
EAYWNGKQGS RPSSSSSYNA NPVQPRLSEN HNLTVTPTYH QNQGENISSL      50
LRELYAKPLS ERHVESDGLI FDPAQITSRT ARGVAVPHGN HYHFIPYEQM     100
SELEKRIARI IPLRYRSNHW VPDSRPEQPS PQSTPEPSPS LQPAPNPQPA     150
PSNPIDEKLV KEAVRKVGDG YVFEENGVSR YIPAKDLSAE TAAGIDSKLA     200
KQESLSHKLG AKKTDLPSSD REFYNKAYDL LARIHQDLLD NKGRQVDFEV     250
LDNLLERLKD VSSDKVKLVD DILAFLAPIR HPERLGKPNA QITYTDDEIQ     300
VAKLAGKYTT EDGYIFDPRD ITSDEGDAYV TPHMTHSHWI KKDSLSEAER     350
AAAQAYAKEK GLTPPSTDHQ DSGNTEAKGA EAIYNRVKAA KKVPLDRMPY     400
NLQYTVEVKN GSLIIPHYDH YHNIKFEWFD EGLYEAPKGY SLEDLLATVK     450
YYVEHPNERP HSDNGFGNAS DHV    (SEQ ID NO : 77)               473
```

FIGURE 42

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN     100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN     150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS     200
ELAAAKAHLA GKNMQPSQLS YSSTASDNNT QSVAKGSTSK PANKSENLQS     250
LLKELYDSPS AQRYSESDGL VFDPAKIISR TPNGVAIPHG DHYHFIPYSK     300
LSALEEKIAR MVPISGTGST VSTNAKPNEV VSSLGSLSSN PSSLTTSKEL     350
SSASDGYIFN PKDIVEETAT AYIVRHGDHF HYIPKSNQIG QPTLPNNSLA     400
TPSPSLPINP GTSHEKHEED GYGFDANRII AEDESGFVMS HGDHNHYFFK     450
KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS HEQDYPGNAK EMKDLDKKIE     500
EKIAGIMKQY GVKRESIVVN KEKNAIIYPH GDHHHADPID EHKPVGIGHS     550
HSNYELFKPE EGVAKKEGNK VYTGEELTNV VNLLKNSTFN NQNFTLANGQ     600
KRVSFSFPPE LEKKLGINML VKLITPDGKV LEKVSGKVFG EGVGNIANFE     650
LDQPYLPGQT FKYTIASKDY PEVSYDGTFT VPTSLAYKMA SQTIFYPFHA     700
GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN AYLENNYKVG EIKLPIPKLN     750
QGTTRTAGNK IPVTFMANAY LDNQSTYIVE   (SEQ ID NO : 78)       780
```

FIGURE 43

```
CAYELGLHQA QTVKENNRVS YIDGKQATQK TENLTPDEVS KREGINAEQI      50
VIKITDQGYV TSHGDHYHYY NGKVPYDAII SEELLMKDPN YQLKDSDIVN     100
EIKGGYVIKV NGKYYVYLKD AAHADNVRTK EEINRQKQEH SQHREGGTSA     150
NDGAVAFARS QGRYTTDDGY IFNASDIIED TGDAYIVPHG DHYHYIPKNE     200
LSASELAAAE AFLSGRENLS NLRTYRRQNS DNTPRTNWVP SVSNPGTTNT     250
NTSNNSNTNS QASQSNDIDS LLKQLYKLPL SQRHVESDGL IFDPAQITSR     300
TARGVAVPHG NHYHFIPYEQ MSELEKRIAR IIPLRYRSNH WVPDSRPEEP     350
SPQPTPEPSP SPQPAPNPQP APSNPIDEKL VKEAVRKVGD GYVFEENGVS     400
RYIPAKNLSA ETAAGIDSKL AKQESLSHKL GAKKTDLPSS DREFYNKAYD     450
LLARIHQDLL DNKGRQVDFE ALDNLLERLK DVSSDKVKLV DDILAFLAPI     500
RHPERLGKPN AQITYTDDEI QVAKLAGKYT TEDGYIFDPR DITSDEGDAY     550
VTPHMTHSHW IKKDSLSEAE RAAAQAYAKE KGLTPPSTDH QDSGNTEAKG     600
AEAIYNRVKA AKKVPLDRMP YNLQYTVEVK NGSLIIPHYD HYHNIKFEWF     650
DEGLYEAPKG YTLEDLLATV KYYVEHPNER PHSDNGFGNA               690
(SEQ ID NO : 79)
```

FIGURE 44

```
GTGAAGAAAA CATATGGTTA TATCGGCTCA GTTGCTGCCA TTTTACTAGC TACTCATATT      60
GGAAGTTACC AACTTGGTAA GCATCATATG GGTCTAGCAA CAAAGGACAA TCAGATTGCC     120
TATATTGATG ACAGCAAAGG TAAGGCAAAA GCCCCTAAAA CAAACAAAAC GATGGATCAA     180
ATCAGTGCTG AAGAAGGCAT CTCTGCTGAA CAGATCGTAG TCAAAATTAC TGACCAAGGC     240
TATGTGACCT CACACGGTGA CCATTATCAT TTTTACAATG GGAAAGTTCC TTATGATGCG     300
ATTATTAGTG AAGAGTTGTT GATGACGGAT CCTAATTACC GTTTTAAACA ATCAGACGTT     360
ATCAATGAAA TCTTAGACGG TTACGTTATT AAAGTCAATG GCAACTATTA TGTTTACCTC     420
AAGCCAGGTA GTAAGCGCAA AAACATTCGA ACCAAACAAC AAATTGCTGA GCAAGTAGCC     480
AAAGGAACTA AAGAAGCTAA AGAAAAAGGT TTAGCTCAAG TGGCCCATCT CAGTAAAGAA     540
GAAGTTGCGG CAGTCAATGA AGCAAAAAGA CAAGGACGCT ATACTACAGA CGATGGCTAT     600
ATTTTTAGTC CGACAGATAT CATTGATGAT TTAGGAGATG CTTATTTAGT ACCTCATGGT     660
AATCACTATC ATTATATTCC TAAAAAGGAT TTGTCTCCAA GTGAGCTAGC TGCTGCACAA     720
GCCTACTGGA GTCAAAAACA AGGTCGAGGT GCTAGACCGT CTGATTACCG CCCGACACCA     780
GCCCCAGGTC GTAGGAAAGC CCCAATTCCT GATGTGACGC CTAACCCTGG ACAAGGTCAT     840
CAGCCAGATA ACGGTGGCTA TCATCCAGCG CCTCCTAGGC CAAATGATGC GTCACAAAAC     900
AAACACCAAA GAGATGAGTT TAAAGGAAAA ACCTTTAAGG AACTTTTAGA TCAACTACAC     960
CGTCTTGATT TGAAATACCG TCATGTGGAA GAAGATGGGT TGATTTTTGA ACCGACTCAA    1020
GTGATCAAAT CAAACGCTTT TGGGTATGTG GTGCCTCATG GAGATCATTA TCATATTATC    1080
CCAAGAAGTC AGTTATCACC TCTTGAAATG GAATTAGCAG ATCGATACTT AGCTGGCCAA    1140
ACTGAGGACA ATGACTCAGG TTCAGAGCAC TCAAAACCAT CAGATAAAGA AGTGACACAT    1200
ACCTTTCTTG GTCATCGCAT CAAAGCTTAC GGAAAAGGCT TAGATGGTAA ACCATATGAT    1260
ACGAGTGATG CTTATGTTTT TAGTAAAGAA TCCATTCATT CAGTGGATAA ATCAGGAGTT    1320
ACAGCTAAAC ACGGAGATCA TTTCCACTAT ATAGGATTTG GAGAACTTGA ACAATATGAG    1380
TTGGATGAGG TCGCTAACTG GGTGAAAGCA AAAGGTCAAG CTGATGAGCT TGCTGCTGCT    1440
TTGGATCAGG AACAAGGCAA AGAAAAACCA CTCTTTGACA CTAAAAAAGT GAGTCGCAAA    1500
GTAACAAAAG ATGGTAAAGT GGGCTATATG ATGCCAAAAG ATGGTAAGGA CTATTTCTAT    1560
GCTCGTGATC AACTTGATTT GACTCAGATT GCCTTTGCCG AACAAGAACT AATGCTTAAA    1620
GATAAGAAGC ATTACCGTTA TGACATTGTT GACACAGGTA TTGAGCCACG ACTTGCTGTA    1680
GATGTGTCAA GTCTGCCGAT GCATGCTGGT AATGCTACTT ACGATACTGG AAGTTCGTTT    1740
GTTATCCCAC ATATTGATCA TATCCATGTC GTTCCGTATT CATGGTTGAC GCGCGATCAG    1800
ATTGCAACAG TCAAGTATGT GATGCAACAC CCCGAAGTTC GTCCGGATGT ATGGTCTAAG    1860
CCAGGGCATG AAGAGTCAGG TTCGGTCATT CCAAATGTTA CGCCTCTTGA TAAACGTGCT    1920
GGTATGCCAA ACTGGCAAAT TATCCATTCT GCTGAAGAAG TTCAAAAAGC CCTAGCAGAA    1980
GGTCGTTTTG CAACACCAGA CGGCTATATT TTCGATCCAC GAGATGTTTT GGCCAAAGAA    2040
ACTTTTGTAT GGAAAGATGG CTCCTTTAGC ATCCCAAGAG CAGATGGCAG TTCATTGAGA    2100
ACCATTAATA AATCTGATCT ATCCCAAGCT GAGTGGCAAC AAGCTCAAGA GTTATTGGCA    2160
AAGAAAAATA CTGGTGATGC TACTGATACG GATAAACCCA AAGAAAAGCA ACAGGCAGAT    2220
AAGAGCAATG AAAACCAACA GCCAAGTGAA GCCAGTAAAG AAGAAAAAGA ATCAGATGAC    2280
TTTATAGACA GTTTACCAGA CTATGGTCTA GATAGAGCAA CCCTAGAAGA TCATATCAAT    2340
CAATTAGCAC AAAAAGCTAA TATCGATCCT AAGTATCTCA TTTTCCAACC AGAAGGTGTC    2400
CAATTTTATA ATAAAAATGG TGAATTGGTA ACTTATGATA TCAAGACACT TCAACAAATA    2460
AACCCTTAA    (SEQ ID NO : 80)                                         2469
```

FIGURE 45

```
VKKTYGYIGS VAAILLATHI GSYQLGKHHM GLATKDNQIA YIDDSKGKAK      50
APKTNKTMDQ ISAEEGISAE QIVVKITDQG YVTSHGDHYH FYNGKVPYDA     100
IISEELLMTD PNYRFKQSDV INEILDGYVI KVNGNYYVYL KPGSKRKNIR     150
TKQQIAEQVA KGTKEAKEKG LAQVAHLSKE EVAAVNEAKR QGRYTTDDGY     200
IFSPTDIIDD LGDAYLVPHG NHYHYIPKKD LSPSELAAAQ AYWSQKQGRG     250
ARPSDYRPTP APGRRKAPIP DVTPNPGQGH QPDNGGYHPA PPRPNDASQN     300
KHQRDEFKGK TFKELLDQLH RLDLKYRHVE EDGLIFEPTQ VIKSNAFGYV     350
VPHGDHYHII PRSQLSPLEM ELADRYLAGQ TEDNDSGSEH SKPSDKEVTH     400
TFLGHRIKAY GKGLDGKPYD TSDAYVFSKE SIHSVDKSGV TAKHGDHFHY     450
IGFGELEQYE LDEVANWVKA KGQADELAAA LDQEQGKEKP LFDTKKVSRK     500
VTKDGKVGYM MPKDGKDYFY ARDQLDLTQI AFAEQELMLK DKKHYRYDIV     550
DTGIEPRLAV DVSSLPMHAG NATYDTGSSF VIPHIDHIHV VPYSWLTRDQ     600
IATVKYVMQH PEVRPDVWSK PGHEESGSVI PNVTPLDKRA GMPNWQIIHS     650
AEEVQKALAE GRFATPDGYI FDPRDVLAKE TFVWKDGSFS IPRADGSSLR     700
TINKSDLSQA EWQQAQELLA KKNTGDATDT DKPKEKQQAD KSNENQQPSE     750
ASKEEKESDD FIDSLPDYGL DRATLEDHIN QLAQKANIDP KYLIFQPEGV     800
QFYNKNGELV TYDIKTLQQI NPP    (SEQ ID NO : 81)              823
```

FIGURE 46

```
GTGAAGAAAA CATATGGTTA TATCGGCTCA GTTGCTGCCA TTTTACTAGC TACTCATATT      60
GGAAGTTACC AACTTGGTAA GCATCATATG GGTCTAGCAA CAAAGGACAA TCAGATTGCC     120
TATATTGATG ATAGCAAAGG TAAGGCAAAA GCCCCTAAAA CAAACAAAAC GATGGATCAA     180
ATCAGTGCTG AAGAAGGCAT CTCTGCTGAA CAGATCGTAG TCAAAATTAC TGACCAAGGT     240
TATGTGACCT CACACGGTGA CCATTATCAT TTTTACAATG GGAAAGTTCC TTATGATGCG     300
ATTATTAGTG AAGAGTTGTT GATGACGGAT CCTAATTACC ATTTTAAACA ATCAGACGTT     360
ATCAATGAAA TCTTAGACGG TTACGTTATT AAAGTCAATG GCAACTATTA TGTTTACCTC     420
AAGCCAGGTA GTAAGCGCAA AAACATTCGA ACCAAACAAC AAATTGCTGA GCAAGTAGCC     480
AAAGGAACTA AAGAAGCTAA AGAAAAAGGT TTAGCTCAAG TGGCCCATCT CAGTAAAGAA     540
GAAGTTGCGG CAGTCAATGA AGCAAAAAGA CAAGGACGCT ATACTACAGA CGATGGCTAT     600
ATTTTTAGTC CGACAGATAT CATTGATGAT TTAGGAGACG CTTATTTAGT ACCTCATGGT     660
AATCACTATC ATTATATTCC TAAAAAAGAT TTGTCTCCAA GTGAGCTAGC TGCTGCACAA     720
GCTTACTGGA GTCAAAAACA AGGTCGAGGT GCTAGACCGT CTGATTACCG CCCGACACCA     780
GCCCCAGGTC GTAGGAAAGC TCCAATTCCT GATGTGACGC CTAACCCTGG ACAAGGTCAT     840
CAGCCAGATA ACGGTGGCTA TCATCCAGCG CCTCCTAGGC CAAATGATGC GTCACAAAAC     900
AAACACCAAA GAGATGAGTT TAAAGGAAAA ACCTTTAAGG AACTTTTAGA TCAACTACAC     960
CGTCTTGATT TGAAATACCG TCATGTGGAA GAAGATGGGT TGATTTTTGA ACCGACTCAA    1020
GTGATCAAAT CAAACGCTTT TGGGTATGTG GTGCCTCATG GAGATCATTA TCATATTATC    1080
CCAAGAAGTC AGTTATCACC TCTTGAAATG GAATTAGCAG ATCGATACTT AGCCGGTCAA    1140
ACTGAGGACA ATGATTCAGG TTCAGATCAC TCAAAACCAT CAGATAAAGA AGTGACACAT    1200
ACCTTTCTTG GTCATCGCAT CAAAGCTTAC GGAAAAGGCT TAGATGGTAA ACCATATGAT    1260
ACGAGTGATG CTTATGTTTT TAGTAAAGAA TCCATTCATT CAGTGGATAA ATCAGGAGTT    1320
ACAGCTAAAC ACGGAGATCA TTTCCACTAT ATAGGATTTG GAGAACTTGA ACAATATGAG    1380
TTGGATGAGG TCGCTAACTG GGTGAAAGCA AAAGGTCAAG CTGATGAGCT TGCTGCTGCT    1440
TTGGATCAGG AACAAGGCAA AGAAAAACCA CTCTTTGACA CTAAAAAAGT GAGTCGCAAA    1500
GTAACAAAAG ATGGTAAAGT GGGCTATATT ATGCCAAAAG ATGGCAAGGA CTATTTCTAT    1560
GCTCGTGATC AACTTGATTT GACTCAGATT GCCTTTGCCG AACAAGAACT AATGCTTAAA    1620
GATAAGAACC ATTACCGTTA TGACATTGTT GACACAGGTA TTGAGCCACG ACTTGCTGTA    1680
GATGTGTCAA GTCTGCCGAT GCATGCTGGT AATGCTACTT ACGATACTGG AAGTTCGTTT    1740
GTTATCCCTC ATATTGATCA TATCCATGTC GTTCCGTATT CATGGTTGAC GCGCGATCAG    1800
ATTGCAACAA TCAAGTATGT GATGCAACAC CCCGAAGTTC GTCCAGATGT ATGGTCTAAG    1860
CCAGGGCATG AAGAGTCAGG TTCGGTCATT CCAAATGTTA CGCCTCTTGA TAAACGTGCT    1920
GGTATGCCAA ATTGGCAAAT CATCCATTCT GCTGAAGAAG TTCAAAAAGC CCTAGCAGAA    1980
GGTCGTTTTG CAACACCAGA CGGCTATATT TTCGATCCAC GAGATGTTTT GGCCAAAGAA    2040
ACTTTTGTAT GGAAAGATGG CTCCTTTAGC ATCCCAAGAG CAGATGGCAG TTCATTGAGA    2100
ACCATTAATA AATCTGATCT ATCCCAAGCT GAGTGGCAAC AAGCTCAAGA GTTATTGGCA    2160
AAGAAAAACG CTGGTGATGC TACTGATACG GATAAACCCA AAGAAAAGCA ACAGGCAGAT    2220
AAGAGCAATG AAAACCAACA GCCAAGTGAA GCCAGTAAAG AAGAAGAAAA AGAATCAGAT    2280
GACTTTATAG ACAGTTTACC AGACTATGGT CTAGATAGAG CAACCCTAGA AGATCATATC    2340
AATCAATTAG CACAAAAAGC TAATATCGAT CCTAAGTATC TCATTTTCCA ACCAGAAGGT    2400
GTCCAATTTT ATAATAAAAA TGGTGAATTA GTAACTTATG ATATCAAGAC GCTTCAACAA    2460
ATAAACCCTT AA    (SEQ ID NO : 82)                                     2472
```

FIGURE 47

```
VKKTYGYIGS VAAILLATHI GSYQLGKHHM GLATKDNQIA YIDDSKGKAK      50
APKTNKTMDQ ISAEEGISAE QIVVKITDQG YVTSHGDHYH FYNGKVPYDA     100
IISEELLMTD PNYHFKQSDV INEILDGYVI KVNGNYYVYL KPGSKRKNIR     150
TKQQIAEQVA KGTKEAKEKG LAQVAHLSKE EVAAVNEAKR QGRYTTDDGY     200
IFSPTDIIDD LGDAYLVPHG NHYHYIPKKD LSPSELAAAQ AYWSQKQGRG     250
ARPSDYRPTP APGRRKAPIP DVTPNPGQGH QPDNGGYHPA PPRPNDASQN     300
KHQRDEFKGK TFKELLDQLH RLDLKYRHVE EDGLIFEPTQ VIKSNAFGYV     350
VPHGDHYHII PRSQLSPLEM ELADRYLAGQ TEDNDSGSDH SKPSDKEVTH     400
TFLGHRIKAY GKGLDGKPYD TSDAYVFSKE SIHSVDKSGV TAKHGDHFHY     450
IGFGELEQYE LDEVANWVKA KGQADELAAA LDQEQGKEKP LFDTKKVSRK     500
VTKDGKVGYI MPKDGKDYFY ARDQLDLTQI AFAEQELMLK DKNHYRYDIV     550
DTGIEPRLAV DVSSLPMHAG NATYDTGSSF VIPHIDHIHV VPYSWLTRDQ     600
IATIKYVMQH PEVRPDVWSK PGHEESGSVI PNVTPLDKRA GMPNWQIIHS     650
AEEVQKALAE GRFATPDGYI FDPRDVLAKE TFVWKDGSFS IPRADGSSLR     700
TINKSDLSQA EWQQAQELLA KKNAGDATDT DKPKEKQQAD KSNENQQPSE     750
ASKEEEKESD DFIDSLPDYG LDRATLEDHI NQLAQKANID PKYLIFQPEG     800
VQFYNKNGEL VTYDIKTLQQ INPP    (SEQ ID NO : 83)             824
```

FIGURE 48

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☑ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

**IMAGES ARE BEST AVAILABLE COPY.**
**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**